BALANCING SAFETY AND DEVELOPMENTAL POTENCY:

ASSESSING THE EFFECT OF ANTIOXIDANTS AND TET ON GENOMIC IMPRINTING STABILITY

A thesis submitted in partial fulfilment

for the degree of Bachelor of Biomedical Sciences with Honours

(Reproduction, Genetics, and Development)

Michael Dunnet

Supervisor: Dr Timothy A. Hore

University of Otago, Dunedin

2018

Abstract

Developmental potency represents the ability of undifferentiated cells to undergo transformation into differentiated cells, with specialised functions. The inner cell mass (ICM) of a blastocyst represents the most developmentally potent cells of the body. The ICM, along with the naïve embryonic stem cells derived from it, possess remarkably low levels of global DNA methylation; a trait that is intrinsic to their pluripotent phenotype and when artificially induced can help re-programme differentiated cells back into a naïve state.

Ten-eleven translocation (TET) enzymes actively remove DNA methylation in a Fe^{2+}-dependent fashion. By assisting Fe^{2+} recycling, the antioxidant ascorbate has been shown to increase TET enzymatic activity and enhance reprogramming of differentiated cells to the naïve state. However, it is currently unclear if TET activation via ascorbate, or other antioxidants, can cause unwanted demethylation at imprint control regions; regulatory elements that control parent-of-origin specific gene expression. Loss of methylation at these regulatory groups results in the biallelic expression of the genes under their control. This phenomenon has been implicated in the formation of a range of cancers, as well as Beckwith-Wiedemann syndrome, a congenital overgrowth disorder.

In this project, I aimed to test what effect the antioxidants ascorbate and hydroquinone have on the stability of imprint control regions through the activation of TET proteins, and in doing so, assess the safety and practicality of using naïve embryonic stem cells for future medical applications. I grew several naïve embryonic stem cell lines, including a TET triple knockout line with inducible TET expression, in varying concentrations of ascorbate or hydroquinone. Bisulfite amplicon sequencing

and florescence activated cell sorting techniques were used to assess methylation patterns at the KCNQ1ot1 imprinted loci in the presence or absence of increased TET activity.

While the study has been hampered by technical difficulties (i.e. clonal amplification was detected, particularly in TET knockout cell lines), I have shown that ascorbate causes significant demethylation of the KCNQ1ot1 imprinted region at 25 ng/μL of ascorbate. While other concentrations, (12.5 ng/μL and 50 ng/μL of ascorbate) were not statistically significant there appeared to be a trend of decreasing methylation which may have become more obvious had the cells been cultured with ascorbate for longer. Interestingly this effect was not seen at 100 ng/μL of ascorbate, indicating there is an optimal concentration range for TET-induced demethylation at the KCNQ1ot1 locus. In contrast, hydroquinone had no significant effect.

Ultimately, these results add to the current literature describing the effects of different culture media constituents on naïve embryonic stem cells. This may allow for the rational design of culture media that can improve both the efficiency and safety of naïve embryonic stem cells and embryos used in regenerative medicine, mammalian transgenics and IVF.

Acknowledgements

I would like to start by expressing my deepest appreciation to my supervisor Dr Timothy A. Hore. Without his persistent guidance and encouragement throughout the year this project would not have been possible. Thank you for all the time and effort you have dedicated to me, it has been an honour to be your student.

I would like to extend my thanks to all the members of the Hore laboratory group. It has been amazing to work alongside all of you. In particular, I would like to thank Dr Donna Bond and Oscar Ortega-Recalde for their help with sample processing and DNA sequencing.

I would also like to thank our research technician, Olga Kardailsky, for her amazing help within the lab, and to Michelle Wilson from the Microbiology & Immunology Department for her help with flow cytometry.

I would like to extend my appreciation to my close friends Adva, Akasha, and Ali. Thank you for all the support and friendship you have given me over the last 4 years. Throughout the good times and bad, you have always had my back. Looking back, my time at university will be an experience I will always cherish.

Lastly, I would like to express my heartfelt gratitude to my family. Thank you for all the support and encouragement you have given me, thank you for all your patience, and thank you for believing in me. I would not be where I am today if not for all of you.

Table of contents

List of Abbreviations

5hmc	5-hydroxymethyl-cytosine
5mc	5-methylcytosine
Asc	Ascorbate
DAPI	4;6-diamino-2-phenylindole
DMR	Differentially methylated region
DMSO	Dimethyl sulfoxide
DNMT	DNA methyltransferase
DNMT TKO	DNA methyltransferase triple knockout
EC	Embryonal carcinoma
EpiSC	Epiblast-like stem cells
ES	Embryonic stem
ESM	Embryonic stem-cell media
FACS	Fluorescence activated cell sorting
FBS	Foetal bovine serum
FGF	Fibroblast growth factor
FPI	Forward priming index
FSC	Side scattered light
FSMQ	Filter sterile Milli-Q
GITC	Guanidinium isothiocyanate
HQ	Hydroquinone
ICM	Inner cell mass
ICR	Imprint control region
iPSC	Induced pluripotent stem cell

IVF	In vitro fertilization
KSR	Knockout serum replacement (media)
LIF	leukaemia inhibitory factor
LOI	Loss of imprinting
mCherry	Monomeric cherry florescent protein
NHSC	Naïve human stem cell
PBAT	Post bisulfite adapter tagging
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
RPI	Reverse priming index
RT-PCR	Reverse transcriptase polymerase chain reaction
SNP	Single nucleotide polymorphism
SPRI	Solid phase reversible immobilization
SSC	Forward scattered light
Tcf3	T-cell factor 3
TET	Ten-eleven translocation
TET TKO	TET triple knockout
TGF	Transforming growth factor
TNA	Total nucleic acid

List of Figures

List of Tables

1 CHAPTER ONE: INTRODUCTION AND AIMS

This introductory chapter describes the relevant literature and rationale for my research undertaken in 2018. Specifically, I will describe how the most developmentally potent cells of an embryo are captured and cultured *in vitro*. In particular, I will focus on a new technique that enables efficient reprogramming back to this naïve state through the use of DNA methylation removal enzymes, known as TETs. While promising, TET-mediated reprogramming has several unexplored safety concerns, particularly around the integrity of genomic imprinting. These safety concerns will form the central point of this thesis.

1.1 Developmental potency

Developmental potency represents the ability of undifferentiated cells to undergo transformation into differentiated cells, with specialised functions. By definition, the first cell of a new organism (i.e. the fertilised egg, or zygote) must be 'totipotent', because it gives rise to all the cells of the embryo and the structures required to support its development. In some mammals (such as cattle and humans), this window of totipotency extends at least to the 2- and 4-cell stage because single cells from these embryos can be isolated and still create a viable organism (Veiga et al., 1987, Johnson et al., 1995).

In contrast to totipotency, pluripotency in mammals is defined as the ability to create all cells of the embryonic body, but not the extraembryonic tissues required for *in utero* development (Martello and Smith, 2014). Cells within the centre of the mouse

blastocyst (i.e. the inner cell mass, or ICM for short) are known to possess pluripotency because they form the embryo proper, the only tissue within the blastocyst that contributes to the developed organism (Martello and Smith, 2014).

To conceptualise the different developmental potential of cells within a growing organism, Conrad Waddington compared development to a hillside landscape. This idea, shown in Figure 1.1, uses a series of ridges and gullies as a metaphor to describe pathways of cell fate, and a marble rolling down any one gully as a representation of the trajectory a cell might take. Once a cell has entered a lineage pathway it is committed, as the ridges prevent it from entering any other gullies (Goldberg et al., 2007, Hemberger et al., 2009). Using this example, totipotent cells are at the top of the landscape, with the potential to move down any gully and form any cell type.

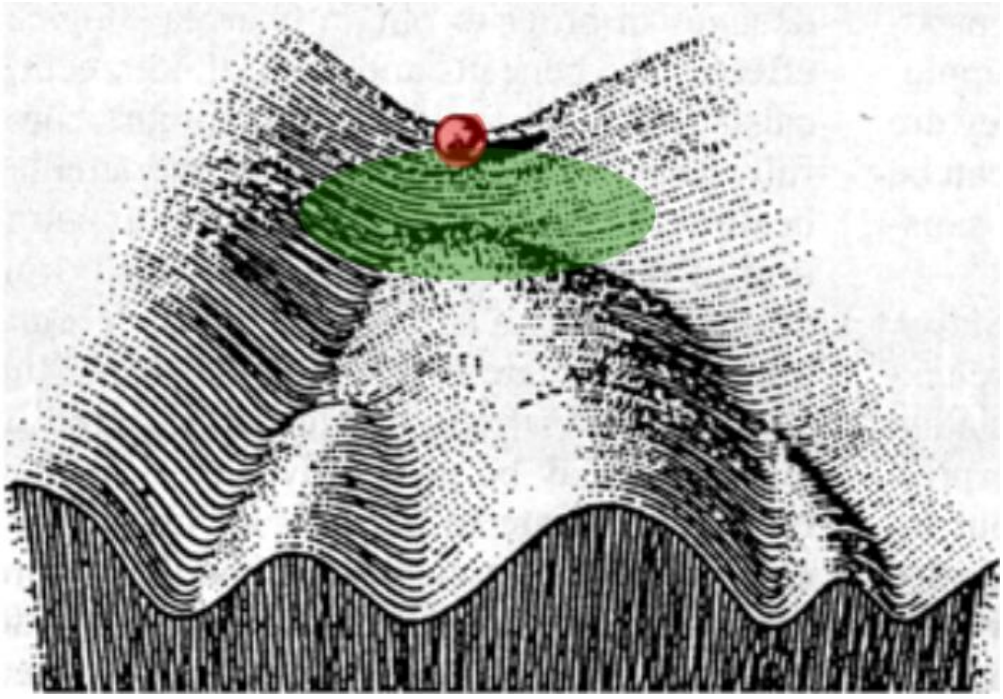


Figure 1.1 Waddington's Epigenetic Landscape.

The marble at the peak of the landscape has several gullies, or cell lineages, it may roll down during development. The ridges prevent the marble from entering adjacent gullies, signalling that once a cell has entered a lineage pathway it has become committed to it, and cannot enter any others. The red circle indicates totipotent cells, while the green area indicates pluripotency.

Pluripotent cells cannot be placed at one point on the landscape but fall anywhere between the top peak and the initial gullies that represent formation into the three germ cell layers. Cells included in this definition can differ in gene expression and epigenetic characteristics (Nichols and Smith, 2009), leading to the idea that pluripotency comes in many 'flavours', with some cells (such as the ICM) considered more pluripotent than others – a concept first proposed by Nichols and Smith, 2009.

1.2 Pluripotent cell culture

The ability to harness the developmental potency of stem cells opens the door for drastic improvements in regenerative medicine or mammalian transgenics. Indeed, multipotent haemopoietic stem cell transplantations are already used to treat many types of leukaemia and lymphoma (Passweg et al., 2012), some genetic diseases such as beta-thalassemia (Morris et al., 2006), and have been described as an effective treatment for HIV/AIDS (Younan et al., 2013, Hutter et al., 2009). However, cells higher up the developmental hierarchy clearly have more flexibility and power; thus, culture of totipotent or pluripotent cells that can be directed to make a variety of tissues or cell-types would be extremely valuable for regenerative medicine. In addition to this, the potential uses of totipotent and pluripotent cells are not limited to a clinical setting. These cell types can offer great insight into the inner workings of mammalian development, furthering our understanding of how a single cell generates an entire organism. While it is currently not possible to culture totipotent cells (Lu and Zhang, 2015), pluripotent cell culture has been employed in developmental biology labs in various forms since the 1950s (Solter, 2006).

This next section will cover the history of pluripotent cell culture from their conception up until the current day, and while doing so will explain the need for naïve pluripotent stem cells as opposed to classically derived 'primed' pluripotent stem cells.

1.2.1 Classical pluripotent cell culture

1.2.1.1 Mice and other Mammals

With mice as the model organism of choice, pluripotent cell culture has been around since the 1950s in the form of embryonal carcinoma (EC) cells. These cell lines, as their name suggests, were derived from carcinomas; more specifically, teratocarcinomas. These tumours contain a mixture of differentiated cell types, including all three germ-layers, which were extracted and cultured *ex vivo* (Solter, 2006). Due to their cancerous origin, these cells often presented with abnormal karyotypes. While this made them inappropriate for clinical applications, they were used as an early tool to study developmental processes.

The culture of pluripotent stem cells was revolutionised in 1981 by Evans and Kaufman, (1981) and Martin, (1981), who independently produced cell lines directly from the pre-implantation embryo of mice. These cultured cells were called embryonic stem (ES) cells, and had a normal karyotype, could be maintained indefinitely, and could easily be injected back into a blastocyst to produce a functional chimeric mouse. For the first time, stable pluripotent cells had been cultured. It was soon recognised that genetic manipulations produced in ES cells could be passed on to adult mice if they were injected into blastocysts and formed the germline of the resulting mouse (Doetschman et al., 1987; Kuehn et al., 1987). This soon led to the creation of the first 'knockout' mouse strain (Koller et al., 1990), which enabled researchers to more

effectively study disease and early developmental processes by inactivating genes of interest and observing the effect, but also allows for the construction of transgenic animals by directly inserting genes into the genome through either viral vector or homologous recombination (Palmiter and Brinster, 1986, Capecchi, 1989). To this day, the vast majority of knockout mice have been produced using ESC technology, culminating in perhaps the most remarkable example of genetic manipulation whereby successive rounds of genetic manipulation have resulted in the replacement of the entire mouse antibody system with that of a human (Lee et al., 2014a).

Despite these initial ground-breaking advancements, ES cell utilisation was hampered by some serious limitations. Firstly, the original ES cell culture system required a fibroblast feeder cell layer, and the media used relied upon undefined (and likely variable) factors present within foetal bovine serum (FBS) (Evans and Kaufman, 1981). While reliance upon feeder cells could be largely circumvented by supplementation of media with leukaemia inhibitory factor (LIF) (Smith et al., 1988, Williams et al., 1988), only in recent years has it been possible to remove reliance upon FBS in culture (Ying et al., 2008). But perhaps most importantly, early ES cell culture was limited because only a portion ES cells grown within classical media conditions contain recognisably pluripotent cells. This was first discovered by analysis of the pluripotency marker NANOG (Chambers et al., 2007), whose expression fluctuates in ES cell culture, with those NANOG negative cells showing some form of differentiation. Later it was shown that other pluripotency-related transcription factors such as REX1, ESRRB and OCT4 show similar heterogeneous expression patterns (Wray et al., 2010). The implication of this is that because some cells lack pluripotency factors, and are apparently ‘primed’ for differentiation, cells cultured in this classical media condition do not possess the complete, or ground state level, pluripotency seen in

the ICM. This heterogeneity is likely due to the undefined culture media, where each cell is experiencing its own unique microenvironment, and therefore varying amounts of pro or anti differentiation signals (Wray et al., 2010). It is possible that this heterogeneity is responsible for perhaps the most significant drawback of the original ES cell culture system; that is, it appears stem cell lines can only be derived in certain situations. For example, while researchers could derive lines from the 129-mouse strain, or related hybrids strains, relatively easily, it was difficult or impossible to do so from other mouse strains, let alone other mammals (Gardner and Brook, 1997). This presented a major barrier to perhaps one of the more useful applications of ES cells, namely mammalian transgenics. If ES cells are ever to be used in this context, they must be produced from animals other than mice.

1.2.1.2 Humans

Perhaps the greatest promise of ES cells, even more so than applications of mammalian transgenics, is regenerative medicine in humans. To do this, human ES cells must be produced. Using classical ES cell culture techniques, Thomson et al., (1998) produced what they described as the first human 'ES' cells using embryos donated from *in vitro* fertilization (IVF) patients. These were grown on mouse embryonic fibroblast feeder cells, with selection and growth of individual colonies resembling human EC cells over several passages. These cells demonstrated pluripotency through teratoma formation upon injection into mice. However, Thompson's ES cells were found to have undergone X-chromosome inactivation, a characteristic associated with the late stage, post-implantation epiblast. Therefore, it is hard to consider these true ES cells, as they represent a much more differentiated stage of development than the already well-established mouse model (Nichols and Smith, 2009).

1.2.2 Naïve ES cells – Culturing the ground state of pluripotency

As previously mentioned, pluripotency can exist in several metastable states. It is clear that mouse and human ES cells grown in regular serum-based media, while displaying many pluripotency-related features, do not represent the most naïve, or ground-state of pluripotency. This is because some or all cells within these cultures display markers of early differentiation such as X-chromosome inactivation or have lost expression of pluripotency-related transcription factors such as NANOG (Nichols and Smith, 2009). Subsequent research efforts focussed upon media conditions which better capture the *in vivo* state of mammalian embryos, specifically the ICM, and have claimed to capture the naïve pluripotency (Ying et al., 2008).

1.2.2.1 Mice

In 2008, ES cells were derived for the first time using a fully-defined media that was not reliant upon serum products such as FBS. The key technological advance allowing this was supplementation with two small molecules, known as '2i', in combination with LIF (Ying et al., 2008). This resulted in homogenous expression of pluripotency markers such as NANOG (Martello and Smith, 2014, Wray et al., 2010) and remarkably similar global DNA methylation patterns to that of the ICM *in vivo* – features not shared with classically grown serum ES cells (Ficz et al., 2013). As such, these cells are considered to possess naive pluripotency, thus earning the name naïve ES cells (Martello and Smith, 2014, Nichols and Smith, 2009, Wray et al., 2010, Ying et al., 2008).

Mechanistically, 2i is comprised of two small molecule inhibitors. Specifically, CHIR99021, which is an inhibitor of GSK3β, and PD0325901, a potent MEK1 inhibitor (Figure 1.2). MEK1 is an integral part of the fibroblast growth factor 4

(FGF4)/ERK pathway in mice (Ying et al., 2008). ES cells naturally express FGF4, which acts in an autocrine manner to induce activation of the ERK/MAPK – a critical pathway required for mesodermal and neural lineage differentiation. Inhibition of this pathway arrests ES cells in a state of naive pluripotency, with full expression of OCT4, NANOG and REX1 (Kunath et al., 2007). OCT4 and NANOG are essential for the formation of the ground state phenotype in naïve ES cells (Nichols and Smith, 2009, Nichols et al., 1998), and while REX1 has no functional advantage for pluripotency attainment, it is exclusively expressed in the pre-implantation embryo and thus a good marker for naïve pluripotency (Kalkan and Smith, 2014). Indeed, antagonism of this pathway is so important for naive ES cell self-renewal that it is considered an integral part of the ground state (Martello and Smith, 2014).

Under normal conditions, GSK3β acts as an inhibitor for the Wnt signalling pathway through phosphorylation of the signal transducing protein, β-catenin. Phosphorylation targets β-catenin for ubiquitination and subsequent destruction (Aberle et al., 1997). GSK3β inhibition, through CHIR99021, enables β-catenin to increase the expression of many pluripotency related genes through the alleviation of a transcriptional repressor protein, T-cell factor 3 (Tcf3) (Wray et al., 2011). Specifically, Esrrb expression is increased, and through interactions with other core pluripotency factors such as NANOG, OCT4, SOX2, and KLF4 (Martello et al., 2012), drastically increases self-renewal and growth, while inhibiting neural differentiation (Ying et al., 2008).

Interestingly, PD0325901 alone leads to progressive degeneration of ES cells, while CHIR99021 alone actively promotes non-neural differentiation. However, when CHIR99021 and PD0325901 are supplemented together the differentiation effect of CHIR99021 is erased by the inhibitory effects of PD0325901 and the promotion of self-

renewal by CHIR99021 prevents degeneration (Ying et al., 2008). Thus, only the combination of both components of 2i can derive naïve pluripotency from classical ES cells.

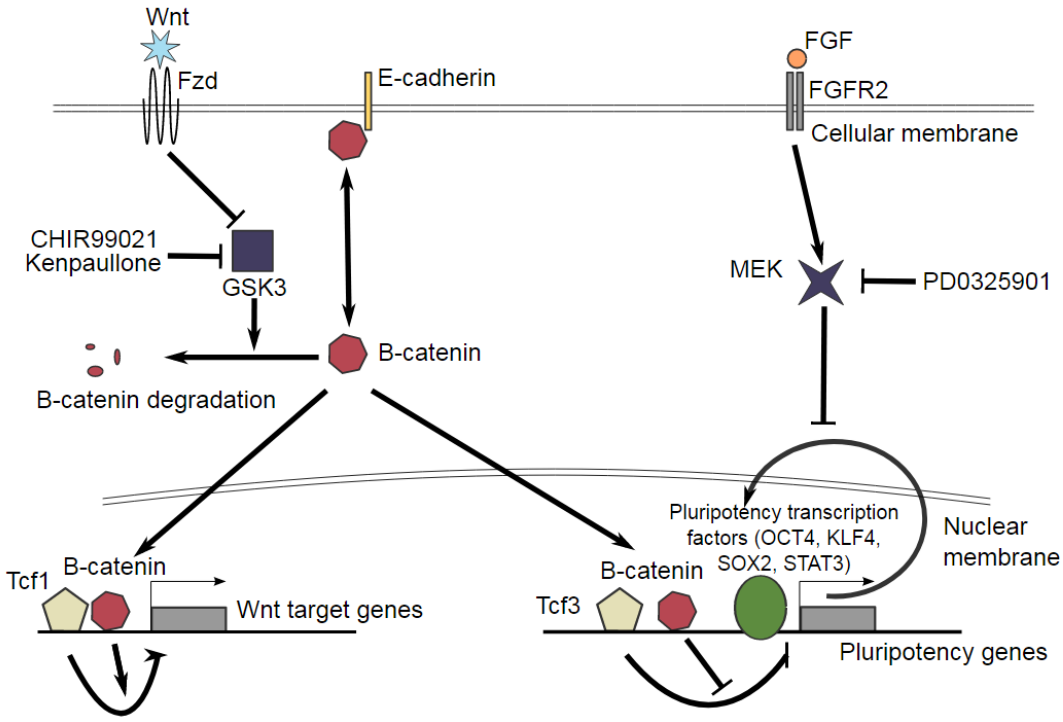


Figure 1.2. 2i signalling

2i, comprised of the two inhibitors CHIR99021 and PD0325901, drives naive pluripotency through the inhibition of GSK3-β and MEK respectively, allowing for the transcription of pluripotency associated genes.

Not only does this discovery help to further the understanding of mammalian development, it has also helped to overcome many of the limitations of classical ES cells. Most importantly, naïve ES cells can be cultured from a wider range of organisms, not just the 129-mouse strain. All strains of mice (Ying et al., 2008), as well as the rat (Li et al., 2008), can be used to obtain naïve ES cells, allowing for full genetic manipulation of both species. This has the potential to greatly improve the scope of biomedical research, as rats have significant benefits over mice as a model organism in many biological investigations (Iannaccone and Jacob, 2009). Further excitement from

the perspective of regenerative medicine is generated by the possibility of creating naïve pluripotent stem cells from human or non-rodent models.

1.2.2.2 Humans

Naïve ES cells have been derived from human blastocysts and classical human ES cells lines (i.e. those first derived by Thomson et al, 1998), although this was not as easy to achieve compared with mouse. Initial attempts focussed upon 2i media, but relied upon forced expression of pluripotency-related transgenes in order to create a naïve-like state. Subsequent attempts to remove the transgenes resulted in successful culture (Gafni et al., 2013, Theunissen et al., 2014), however, this requires a combination of factors in addition to standard 2i media. For example, Naïve Human Stem Cell (NHSC) media developed by the Hanna laboratory contains 2i as well as FGF-2, TGF-1β, a p38 inhibitor, and JNK inhibitor. Interestingly, whilst FGF-2 is required for the maintenance of the naïve state in humans (Gafni et al., 2013), in mice FGF-2 induces differentiation via activation of the ERK pathway (Ying et al., 2008). This highlights that despite their similarities, naïve ES cells in mice and humans have important biological differences that must be considered when using mice ES cells as a model for human health or disease.

Despite the recent derivation of naïve human ES cells, they are already showing great potential as an option for regenerative medicine. Recently, human ES cell lines have shown promise in treating wet age-related macular degeneration, a condition in which leaky blood vessels grow underneath the retina and excess fluid damages pigment epithelium, disrupting vison (Liu et al., 2018).

1.2.3 Reprogramming to the ground state

In humans, and mice, forced expression of four key pluripotency related transcription factors (c-myc, SOX2, Oct4, and Klf4) can produce induced pluripotent stem cells (iPSCs) from differentiated adult cells (Takahashi and Yamanaka, 2006). iPSCs have several benefits over ES cells for regenerative medicine, including a reduced chance of tissue rejection because the cells are derived directly from the patient requiring treatment (Takahashi and Yamanaka, 2006). Ethical considerations are also lessened with iPSCs when compared to ES cells. This comes from the fact that ES cell lines must be derived from an otherwise normally developing embryo, destroying it in the process (Evans and Kaufman, 1981). Jung (2009) describes the current debate on whether or not human ES cells can be classified as a person, and thus whether or not we have the right to destroy these embryos to generate ES cell lines. It is unlikely that this debate will be resolved in the near future, meaning there is the potential for individuals to object to ES cell therapies based upon their own moral viewpoints. Thus, a focus on reprogramming cells back to a naive state of pluripotency may be required to help ensure the best health outcomes, both from a social and biological standpoint.

While iPSCs represent an exciting prospect with respect to regenerative, as a model system they are somewhat difficult to work with. Principally, it can take over 1 month to reprogram differentiated cells to the pluripotent state. Epiblast-like stem cells (EpiSC) resemble the post-implantation gastrulating embryo and can be reprogrammed to induced pluripotent stem cells (iPSCs) within just a few days following forced expression of at least one component of the core pluripotency circuitry network (e.g. Nanog, Klf4, Klf2, or Esrrb (Martello and Smith, 2014)) and 2i conditions. Yet, this process is inefficient, with reprogramming rates below 1% (Guo et al., 2009). As such, they provide a useful tool to show iPSC derivation can be accelerated or improved.

1.3 Epigenetics and ground state pluripotency

1.3.1 Naïve ES cells and the ICM

Whole genome bisulfite sequencing is a useful tool that has been employed to assess global methylation patterns of both naïve and classical ES cells, as well as early blastocysts and differentiated adult tissues. This has revealed perhaps the most striking feature of naive ES cells: remarkably low global levels of DNA methylation. Only 25-35% CpG dinucleotides throughout the genome are methylated (Ficz et al., 2013, Habibi et al., 2013). This contrasts with classically grown ES cells and adult cell lineages, which have ~75-80% global CpG methylation (Ficz et al., 2013, Habibi et al., 2013, Hon et al., 2013). Furthermore, the pre-implantation ICM of a developing embryo *in vivo* shares the globally low DNA methylation pattern of naive ES cells (Smallwood et al., 2011), suggesting the two might be equivalent. This loss of methylation is characteristic of early mammalian embryos and is thought to 'wipe the slate clean' of epigenetic memory before building it up again as the embryo develops. This prevents acquired epigenetic marks from parental DNA from impacting the offspring (Lee et al., 2014b). Some regions are exempt from demethylation, such as imprint control regions (ICR), as proper methylation patterns here are extremely important for the growth and development of the embryo (Morison et al., 2005).

1.3.2 Demethylation as a driver for naïve pluripotency

Not only is low methylation correlated with naïve pluripotency, but it is also a driver of it. DNA methyltransferase triple knockout (DNMT-TKO) naïve ES cells are deficient in all three methyltransferase genes and cannot add methylation to CpG dinucleotides, and as a result, are dramatically hypomethylated. DNMT-TKO cells are resistant to

differentiation and retain expression of many pluripotency factors even after induction of differentiation (Schmidt et al., 2012). In some cases, differentiation can occur, but when treated with LIF, they readily revert back to the naive state (Schmidt et al., 2012). Consistent with these results is that DNMT3A and DNMT3B, along with their co-factor DNMT3L, are significantly low expression in 2i treatment, but upon switching to the classical serum culture method, are drastically upregulated (Habibi et al., 2013) – a phenomenon attributed to MEK inhibition by PD0325901 (Choi et al., 2017). Because these enzymes are responsible for de novo methylation (Ficz et al., 2013, Habibi et al., 2013, Lee et al., 2014b), their low expression in 2i suggests that hypomethylation in the ground state, is at least in part, due to the inability of naïve ES cells to deposit new methylation marks. DNMT1 is responsible for maintenance of DNA methylation patterns and does so via recruitment to hemi-methylated loci by UHRF1 (Lee et al., 2014b). While protein levels of DNMT1 and UHRF1 are not altered, they cannot sustain DNA methylation on their own in naïve ES cells. As a result, methylation is passively lost when it is not replaced after cell replication (Lee et al., 2014b, Habibi et al., 2013). Together, these data suggest that methylation is required for differentiation.

TET family proteins are also suspected to be involved in the demethylation of naïve ES cells (Ficz et al., 2013). All TET proteins (TET1, TET2, and TET3) are α-ketoglutarate and Fe(II) dependant dioxygenases. They have the ability to remove DNA methylation by oxidation of 5-methyl-cytosine (5mC) to 5-hydroxymethyl-cytosine (5hmC) (Ito et al., 2010, Tahiliani et al., 2009), shown in Figure 1.3a. In addition to this, TET1 has been shown to maintain ES cell self-renewal by upregulating NANOG expression though hypomethylation of its promotor region (Ito et al., 2010). Remarkably, both TET1 and TET2 also share DNA binding sites with NANOG and are recruited to these in a NANOG dependant manner (Costa et al., 2013). Many of these

loci are the promotor regions of pluripotency associated factors, the most prominent of these being Esrrb. This suggests a positive feedback loop between NANOG and TET, resulting in demethylation and expression of pluripotency factors, pushing the cell towards a pluripotent state (Costa et al., 2013). Indeed, culturing ES cells in 2i results in an upregulation of TET2 (Ficz et al., 2013), which could help drive this effect. Contrasting this, loss of proper TET activity within naïve ES cells drastically reduces their viability (Dawlaty et al., 2014), further supporting this idea.

Therefore, the hypomethylation phenotype observed in naive ES cells is likely due to a combination of passive demethylation though failed maintenance after replication, along with active oxidation of 5mC by TET1/2, and the inability to deposit new methylation marks (Ficz et al., 2013). However, more importantly, the interactions between TET proteins and NANOG indicates that not only is pluripotency reliant on hypomethylation, but that hypomethylation is heavily modulated by the components of the pluripotency network (Costa et al., 2013, Lee et al., 2014b). This suggests that globally low DNA methylation and naive pluripotency are profoundly intertwined and that without this core feature, the ground state is difficult to reach.

1.3.3 Inducing demethylation to drive naïve pluripotency

It is clear that global hypomethylation and naïve pluripotency in ES cells are intrinsically tied, leading to the idea that actively inducing hypomethylation may be a pathway towards pluripotency (Hore et al., 2016). As stated above, reprogramming EpiSCs to a naive state of pluripotency is possible, however, the conversion rate is very low at less than 1% (Guo et al., 2009, Martello and Smith, 2014). Recently, increasing the expression and activity of the TET enzymes with two commonly available vitamins,

ascorbate (vitamin C) and retinol (vitamin A), has led to a significant increase in reprogramming efficiency (Hore et al., 2016).

Vitamin C was discovered to enhance TET activity in 2013, however, it was thought to act as a bound co-factor to TET enzymes (Blaschke et al., 2013). This is not the case, as Hore et al, (2016) have shown that ascorbate acts as an antioxidant, reducing Fe(III) back to Fe(II), and indirectly boosting TET activity (Figure 1.3b). As such, any antioxidant which is capable of reducing Fe(III) can be used to enhance TET. From a range of tested antioxidants, vitamin C is the only one to have an impact without a prior incubation period *in vitro,* but many other antioxidants did prove to have an effect (Hore et al., 2016). Interestingly, hydroquinone, the antioxidant with the second largest effect reported by Hore et al., (2016) has also been reported to increase demethylation through TET enzyme activation in cancer cells (Coulter et al., 2013), providing further evidence that the presence of antioxidants enables TET-mediated demethylation.

This advancement has the potential to drastically improve stem cell science and regenerative medicine, as reprogramming human cells to iPSCs may become much more efficient with the addition of antioxidants to culture media. Antioxidant supplementation may also result in reduced reliance on pluripotency factor overexpression when producing iPSCs, which may help reduce the abnormal and potentially unsafe methylation patterns described by (Planello et al., 2014).

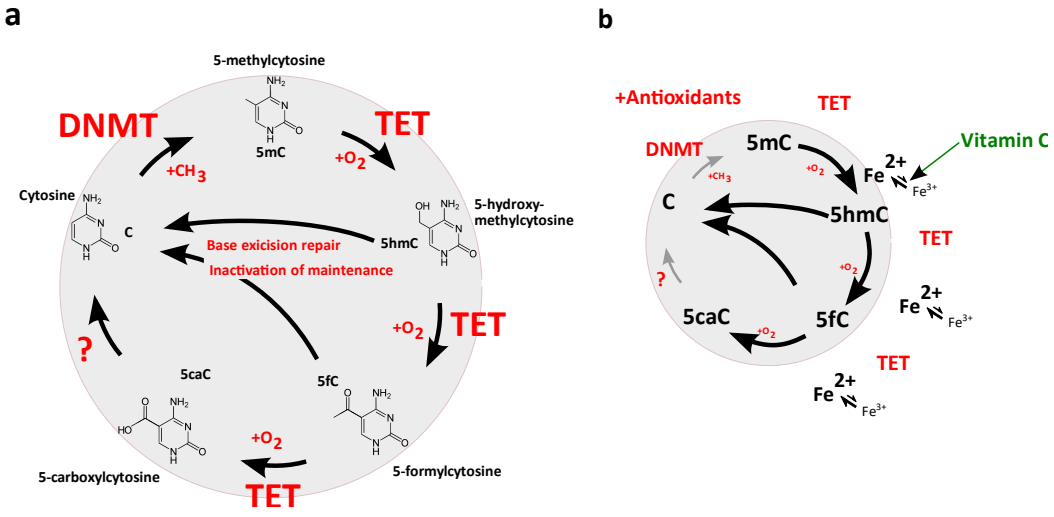


Figure 1.3 Enzymatic activity of TET proteins

Adapted from Hore, T. A. (2017). (**a**)TET proteins remove DNA methylation by oxidation of 5-methyl-cytosine (5mc) to 5-hydroxymethyl-cytosine (5hcm). Base excision repair of 5hcm after TET activity results in conversion back to cytosine. (**b)** Antioxidant treatment enhances Fe^{2+} recycling, resulting in increased TET activity and subsequent demethylation.

1.3.3.1 Potential impacts on genomic imprinting stability

However, actively inducing hypomethylation has several associated risks that have not been fully explored – namely dysregulation at ICRs; cis-acting regulatory elements that control parent-of-origin specific gene expression. ICRs employ differential DNA methylation patterns between alleles to control which parental copy is expressed. This results in monoallelic gene expression in embryos, and later, in the adult. These imprints are established during gametogenesis (Morison et al., 2005), and while there appears to be natural lability in imprinted gene methylation in the cells of the ICM (Lee et al., 2018), methylation-dependant imprinting is maintained throughout an organism's lifespan. However, recent evidence suggests that if severe hypomethylation occurs during this early labile stage, loss of imprinting cannot be restored (Lee et al., 2018).

Loss of imprinting (LOI) is often found at a high frequency in at least 14 different cancers (Jelinic and Shaw, 2007). More specifically, LOI is seen in 100% of chronic myeloid leukaemia cancers, and over 80% of all ovarian cancers (Jelinic and Shaw, 2007), signifying that epigenetic dysregulation is a prime driver of these two cancers. Almost all imprinted genes identified to date are either growth factors or their receptors (Jelinic and Shaw, 2007, Morison et al., 2005), suggesting that hypomethylation at these regions may result in an increased growth rate in affected cells. Thus, it is imperative that the stability of ICRs in ES cells is understood during re-programming with enhanced TET activity, as any cells with affected imprinting will be unsafe for regenerative medicine in humans and ineffective for mammalian transgenics.

While there is some evidence that ICRs are protected from antioxidant induced TET activation, this has only been studied using ascorbate supplementation at a static concentration (Blaschke et al., 2013), but not with other antioxidants such as hydroquinone. In addition, prolonged MEK inhibition, as a result of 2i, can cause aberrant expression of imprinted genes and genome instability (Choi et al., 2017). Therefore, it is hypothesised that TET-induced hypomethylation may exacerbate this instability. Indeed, recent evidence shows that the combination of 2i and a serum free culture media containing vitamin C, known as KSR, drastically increases demethylation at ICRs when compared to 2i + FBS media (Lee et al., 2018).

1.4 Research aims

My research aimed to further investigate if antioxidant-induced TET activation can disrupt the stability of ICRs and result in abnormal biallelic expression of imprinted genes in ES cells. I hypothesised that by steadily increasing the concentration of

antioxidants such as ascorbate or hydroquinone within culture media, mouse ES cells would subsequently lose DNA methylation at imprint control regions in a dose dependant manor.

To test this, I attempted to enhance TET activity by increasing Fe^{2+} recycling with various antioxidant treatments. Specifically, by using LIF + 2i conditions, I cultured the naïve ES cell line B6xSD7, as well as the TET triple knockout (TET TKO) cell line, in increasing amounts of ascorbate or hydroquinone before examining the methylation patterns at the KCNQ1ot1 locus.

2 CHAPTER TWO: MATERIALS & METHODS

2.1 Cultured cells

2.1.1 Mouse embryonic stem cells

Four distinct mouse ESC lines were cultured for experimentation, two wild type and two knock out. The first, B6xSD7, is a wild type cell line derived from the blastocyst of a cross between C57BL6/JOlaHsd and a related congenic strain where the distal arm of chromosome 7 is from SD7. The original motivation behind the derivation of this cell line was to enable investigation of differentially methylated regions (DMRs). One such region is the KCNQ1ot1 locus, which in B6xSD7 harbours an 18 base pair insertion in the maternal (C57BL/6), but not paternal (SD7), allele (Figure 2.1a) (Krueger et al., 2012). V6.5 is another wild type cell line employed in these experiments, derived from the inner cell mass of a C57BL/6 x 129/sv cross (Rideout et al., 2000). C57BL/6 and 129/sv mouse strains differ at two single nucleotide positions in exons two and three of the imprinted gene *Impact* (Choi et al., 2017), allowing for identification of either the maternal (C57BL/6) or paternal (129/sv) allele (Figure 2.1b-c). Two TET TKO cell lines were used for experimentation. TET TKO cell lines are derived from a C57BL/6 and 129/sv background and are deficient in all three TET enzymes, preventing any TET mediated demethylation. Specifically, this line originates from a C57BL/6 x 129/sv Tet1 knock out (KO) crossed with a 129/sv Tet3 KO line, producing a Tet1, Tet3 double knockout (DKO). Targeted Tet2 KO was then performed on the DKO line, producing the TET TKO line. All Tet genes contain missense mutations in their respective catalytic domains (Gu et al., 2011, Hu et al., 2014). A derivative of the TET TKO line was employed, TET TKO mTET1-CD. This line is

transfected with a catalytically active form of TET1, as well as the red fluorescence protein mCherry, under the control of a doxycycline dependant Tet-on inducible promotor.

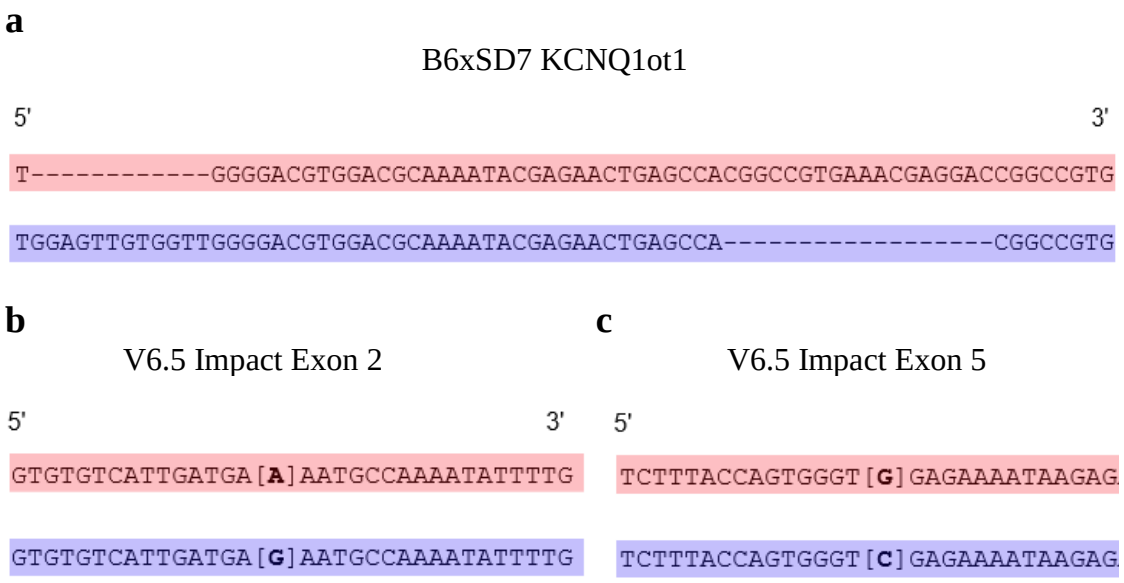


Figure 2.1 Allelic differences at targeted imprinted loci.

(a) Section of sequence from the KCNQ1ot1 imprinted region showing a deletion and insertion in the maternal allele (C57BL6) (red) when compared to the paternal allele (SD7) (blue). **(b)** Section of sequence from the imprinted region *Impact* showing a SNP between the maternal allele (C57BL6) (red) and paternal allele (blue) (129/sv) in both exon 2 (left) and **(c)** exon 5 (right).

2.1.2 Cell culture, passaging, and harvesting

Mouse embryonic stem cells were grown in either 24, 12, or 6-well plates, or 90mm dishes (Culture dishes from Corning®) coated with 0.1% gelatin (G1890-1KG) and cultured with one of two Embryonic Stem cell Media (ESM) variations. The first is an FBS-based ESM supplemented with LIF either with or without 2i (serum + LIF or serum + 2i/LIF). The second is a chemically defined media consisting of N2 (Gibco 17502-048) and B27 (Gibco 12587-101) supplements, as well as LIF and 2i (N2B27 +

2i/LIF). Recipes for both ESM variants can be found in supplementary Table 1 and 2. The 2i mixture can be found in supplementary Table 3.

To prevent overcrowding, cultures were passaged once every two to three days. During passaging, cells were washed twice with phosphate buffered saline (PBS) to remove detached, dead cells and any remaining media. Next, cells were incubated with StemPro® Accutase® (A11105-01) (volume dependant on size of the plate, see supplementary Table 4) for five minutes at 37°C to dissociate all remaining colonies from the plate. Cells were collected in 15 mL falcon tubes containing 10mL PBS and mixed via pipetting up and down to further detach the cells from one another, before being pelleted by centrifugation at 10 RFC for five minutes. Supernatant was then aspirated, and cells were resuspended in the appropriate volume of ESM (supplementary Table 4) before subsequent re-plating.

Upon completion of experiments, media was aspirated from wild type cell lines (B6xSD7 and V6.5), before being washed twice with PBS and lysed in 150uL of Guanidinium Isothiocyanate (GITC). Cell lysates were stored in the -20°C freezer until further use. TET TKO cell lines were harvested post fluorescence activated cell sorting (FACS).

2.1.3 Cell treatments

During ascorbate and hydroquinone variation experiments all cultured cells were grown in the presence of 2i. This results in the rapid genome-wide demethylation of ESCs (Habibi et al., 2013, Lee et al., 2014b), however DNA methylation at imprinted regions is initially protected over the first few passages, before also being lost (Choi et al., 2017). Therefore, there is a time window where genome wide demethylation has occurred, but imprinted region methylation is intact. To allow global

demethylation to occur, but to ensure imprint stability is maintained, all cultures were pre-incubated in 2i conditions for 168 hours (one week) before treatments began.

It was imperative that all experiments were seeded with the same number of cells prior to ascorbate and hydroquinone treatment. This was performed via cell counts with a Neubauer haemocytometer. Firstly, cells were dissociated with Accutase, resuspended in PBS, and pelleted as described above. The supernatant was removed, and cells were once again resuspended in 10mL of PBS. After further mixing via pipetting, 15 μL of cell suspension was examined under a light microscope with a haemocytometer. All cells within the four large corner squares (0.1μL per square) were counted and averaged. This enabled an estimate of total cells in suspension, allowing accurate calculations in seeding densities (supplementary Table 5).

Lastly, TET TKO mTET1-CD cultures were treated with 1 μM doxycycline 72 and 24 hours prior to harvesting by FACS, respectively.

2.1.4 Thawing and freezing cells

Prior to use, all cell lines were stored in liquid nitrogen. All frozen cells were stored in a freezing buffer consisting of 75% serum + 2i/LIF ESM, 15% FBS, and 10% Dimethyl sulfoxide (DMSO). To thaw them, cells were removed from liquid nitrogen and quickly warmed to 37°C. Once thawed, cells were transferred to a 15mL falcon tube containing 10mL of PBS. The suspension was then pelleted, and supernatant aspirated. The appropriate volume of ESM was added depending on plate size (Supplementary Table 6), and cells resuspended. The suspension was then plated out onto cell culture plates coated with 0.1% gelatin.

Cells were regularly frozen down to be used for later experiments. Once the cultures became confluent, cells were collected as described in section *2.1.2 Cell*

culture, passaging and harvesting. Cell counts with a Neubauer haemocytometer (performed as described above in *section 2.1.3 experimental preparation*) were used to estimate cell concentration. One million cells were resuspended in 250 µL of freezing buffer before being transferred to Corning™ 1 mL Cryogenic Storage Vials and stored in the -80°C freezer until further use.

2.2 Total nucleic acid extraction

Both DNA and RNA from cultured cells, collectively referred to as total nucleic acid (TNA), were extracted using the Bio-on magnetic bead approach (Oberacker et al., 2018). For each sample, 60 µL cell lysate, dissolved in GITC, was mixed with 40 µL TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA) and diluted magnetic beads until homogenous. 80 µL of isopropanol was added, and samples were again mixed until homogenous, before being incubated at room temperature for one minute allowing nucleic acids to precipitate and bind to the magnetic beads. Samples were then placed on a magnetic rack for 5 minutes, allowing the beads to accumulate on the side of the tubes and the supernatant was discarded. To remove residual GITC, each sample was washed with 150 µL of isopropanol followed by two subsequent washes with 200 µL of 70% ethanol. 70% ethanol was removed by allowing the beads to dry for 5 minutes. TNA was eluted by removing the samples from the magnetic stand and resuspending the beads in 50 µL of filter sterile Milli-Q (FSMQ) water. This was followed by an incubation at room temperature for one minute, before reapplying the samples to the magnetic, and collecting the supernatant.

2.2.1 Bisulfite conversion

Bisulfite conversion of DNA is a useful tool that enables the visualisation of DNA methylation at single base pair resolution. In short, bisulfite treatment results in deamination of cytosine residues, and subsequent transformation into uracil. However, cytosine methylation renders it resistant to conversion. During PCR, uracil is read as, and subsequently replaced by, thymine and throughout amplification. After sequencing, comparisons back to an untreated sequence are made to assess DNA methylation levels - only methylated cytosines remain following bisulfite treatment, while unmethylated cytosine will appear as thymine.

Once TNA had been isolated, bisulfite conversion was carried out using the Zymo Research EZ-96 DNA Methylation-Direct kit (4x96rxn, D5044) according to the protocol provided. Following conversion, samples were used in either post bisulfite adaptor tagging (PBAT), or PCR to generate amplicon libraries.

2.2.2 cDNA library generation

First strand cDNA synthesis was performed by using the Thermo Scientific RevertAid Reverse Transcriptase (EP0441) kit as per the protocol provided. First strand synthesis was performed using random hexamer primers. Resulting cDNA was amplified using PCR to generate amplicon libraries.

2.3 Primer sets

All primers used had been designed for previous experimentation and were re-used for the current experiments. Some primer sets contain the Otago index handle (Table 4), which are used to attach unique indexes for multiplex sequencing. All primer sets with this feature have the prefix ‘tag’ in their name.

The primer set for the KCNQ1ot1 locus were originally used by Kreuger .,et al (2013), but contain the Otago indexing handle at the 5’ end (Table 2.1). Primers used for *Impact* exon 2 and exon 5 cDNA sequence amplification span the intron exon boundary for their respective exon and were previously employed in the Hore lab at the University of Otago. These primers also contain the Otago indexing handle at the 5’ end (Table 2.2). The primer sets used for *Impact* exon 2 and exon 5 genomic DNA amplification have been previously employed in the Hore lab at the University of Otago and contain no indexing handle (Table 2.3).

Table 2.1 Tagged primers used for KCNQ1ot1 sequence amplification

Primer Name	Sequence
tag_KCNQ-bis-F	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGTTTTAAGATTATTTTTGTTTTTGTAA
tag_KCNQ-bis-R	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTTTCTATTCAACTTAATTCCCAAC

Table 2.2 Tagged primers used for Impact exon 2 and exon 5 cDNA sequence amplification.

Primer Name	Sequence
tag_iSNP_ex2_cF1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGCTGAAGAGGAAGTAGGGAAC
tag_iSNP_ex2_cR1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGATAAGAAGGCGGTGCTGTA
tag_iSNP_ex5_cF1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTACAGCACCGCCTTCTTATCA
tag_iSNP_ex5_cR1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCTTCTTCGGACTCAACCTCAAC

Table 2.3 Primers used for Impact exon 2 & exon 5 genomic DNA seqeunce amplification.

Primer Name	Sequence
iSNP_ex2_gF1	ACAGGTTGAATGTTTATTTGTTGACT
iSNP_ex2_gR1	TCTAATGTCTGGAAAATAGCTGCA
iSNP_ex5_gF1	AACGTCTTGAATTTGTGTTGTTACA
iSNP_ex5_gR1	AAAGAGCACAAAGAGGTACCACT

Table 2.4 Otago indexing handle sequence

Name	Sequence
Otago forward indexing handle	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Otago reverse indexing handle	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

Primers were suspended in the appropriate amount of TE buffer to produce a 200 µM stock solution. A 20 µM working solution was produced by diluting the stock solution by a ratio of 1:10 with FSMQ water. All primers were stored in the -20°C freezer and thawed upon use.

2.4 Polymerase chain reaction

PCR was used to amplify targeted regions of interest for amplicon multiplex sequencing and Sanger sequencing. For multiplex sequencing, unique indexes must be added to each individual sample. This was performed using two successive rounds of PCR and required primers with the additional index handle at the 5' end (Table 4). The first round of PCR used sequence specific primers to amplify target regions, while the second round used primers containing a unique six base pair index which could be used

to identify specific samples post sequencing. Figure 2.2 shows a schematic diagram of the two step PCR process. Indexing primer sequences can be found in supplementary Table 7 and 8. To prepare samples for Sanger sequencing a single round of PCR was performed using primers without the index handle.

PCR optimization was required for successful amplification of target regions. This included gradient PCR tests to determine the most optimal annealing temperature, changes in primer and template concentration to reduce primer dimer if any was present, changes to the elongation time to increase the yield of longer PCR products, and changes in cycle number to adequately attach sequencing indexes when appropriate. Specific changes are outlined in *Chapter 3: Results* when relevant, however, optimisation resulted in the PCR parameters outlined below.

2.4.1 Thermal cycling parameters

Bisulfite converted DNA was amplified using the KAPA Taq ReadyMix PCR Kit (KR0354 – v7.16) as per the manufacturer's instructions. The components of the PCR are listed in Table 2.6. All other PCRs were performed with the Phusion® High-Fidelity (HF) DNA Polymerase (NEB, M0530L), following manufacturer's instructions. The components of this reaction are listed in Table 2.5.

The KCNQ1ot1 locus amplification Thermal cycling was carried out using the following cycle parameters:

92C, 2min; 25(92C, 10s; 59C, 10s; 72C, 30s); 72C, 5min;* ***4C, infinite****; 10**(92C, 10s; 59C, 10s; 72C, 30s); 72C, 5min; 4C, infinite.*

The bolded step indicates a pause where indexes were added.

To amplify exon 2 and exon 5 from the *Impact* locus for Sanger sequencing the following thermal cycling parameters were used:

98°C, 2 min; 34*(98°C, 10 sec; 65°C, 10 sec; 72°C, 30 sec); 4°C ∞.

When amplifying the same regions for multiplex sequencing the cycling parameters were changed to the following:

98°C, 2 min; 28(98°C, 10 sec;* 65°C*, 10 sec; 72°C, 30 sec);* ***4°C*** *∞****;*** *5*(98°C, 10 sec;* 65°C, *10 sec; 72°C, 30 sec); 8 ∞.*

Bolded 4°C ∞ step is the hold stage where indexes are added.

Table 2.5 KAPA polymerase chain reaction components

KAPA		
Reagent	**Final conc**	**10uL reaction**
KAPA (2X)	-	5
Primer (Forward)	0.3 uM	0.15
Primer (Reverse)	0.3 uM	0.15
Template	-	0.5
Milli-Q H20	50-X	7.2
Total	-	10
Index concentrations		
Forward priming index (FPI)	0.2 uM	0.2
Reverse priming index (RPI)	0.2 uM	0.2

Table 2.6 Phusion polymerase chain reaction components

Phusion (Finnzyme/NEB)		
Reagent	**Final conc**	**10uL reaction**
Buffer (HF buffer)	1X	2
dNTPs (stock with 5mM each) (BIO-39025)	200uM each	0.4
Primer (Forward)	0.5uM	0.25
Primer (Reverse)	0.5uM	0.25
Template	-	0.4
Taq (Phusion)	0.02uM	0.1
H20	50-X	6.8
Total	50	10
Index concentrations		
Forward priming index (FPI)	0.2 uM	0.2
Reverse priming index (RPI)	0.2 uM	0.2

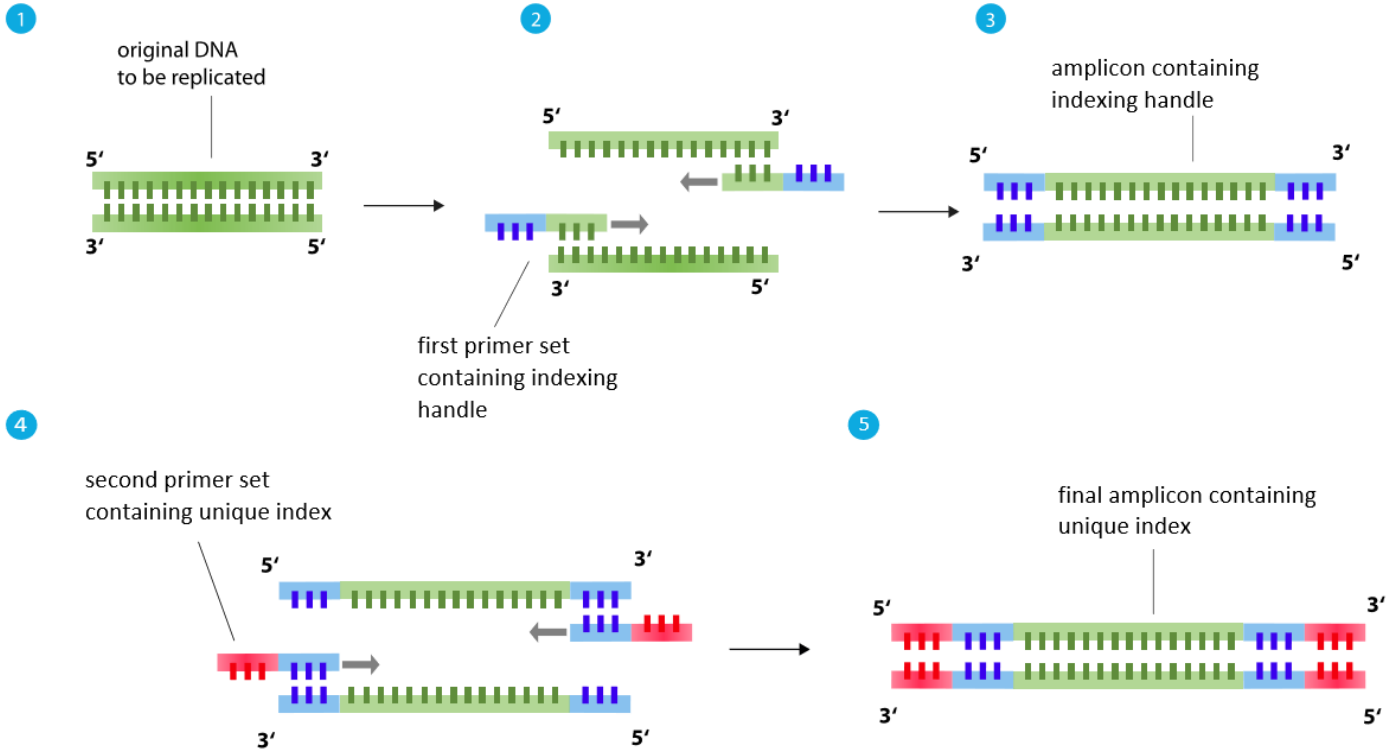


Figure 2.2 Addition of unique indexes PCR amplicons.

To attach unique indexes to each strand of template DNA two successive rounds of PCR were performed. In the first round, primers containing both sequence complementary to the target strand (green), as well as an indexing ‘handle’ (blue) were used (2). Upon completion, the indexing handle has effectively been added to both the 5’ and 3’ end of the template amplicon. The second round uses primers containing sequence complementary to the indexing handle (blue), as well as a unique index which is different for both forward and reverse primers (red). Once the second cycle is complete, unique indexes have been attached to the original strand, resulting in amplicons that can be multiplex sequenced.

2.5 Agarose gel electrophoresis and extraction

To analyse DNA and RNA integrity after extraction, and to visualise PCR product fragment size after amplification, agarose gel electrophoresis was employed. Agarose gel was prepared at 1% weight/volume in Tris/ Acetate/ EDTA (TAE) buffer. GelRed (Biotium 41003) was added at a ratio of 1/20,000 in order visualise DNA and RNA fragments under ultra violet (UV) light. PCR products were loaded into wells along with the 100-base pair Bioline HyperLadder™ (Supplementary Figure 6) and run at 110V for 60 minutes. A UV light transilluminator was used to visualise and photograph DNA fragments.

2.5.1 Gel extraction

Gel extraction was carried out when purifying amplicons for multiplex sequencing in the cases where an excess of primer/index dimer was present. Agarose gels were prepared as described above, however, GelGreen (Biotium 41005) dye was used in place of GelRed at a concentration of 1/20,000. All samples for sequencing were pooled together and run in a single lane alongside a DNA ladder. Using a Midi LED Transilluminator, which can visualise GelGreen using blue light, fragments were excised from the gel using a scalpel. DNA extraction from the gel was carried out using the QIAquick Gel Extraction Kit protocol (Qiagen, cat. 28704) as per the manufacturer's instructions. DNA was eluted in 30 μL of Buffer EB (10 mM Tris-HCl, pH 8.5) and concentration was assessed using a nanodrop spectrometer.

2.6 Florescence activated cell sorting

The TET TKO mTET1-CD cell line possesses a doxycycline dependant Tet-on inducible promotor which controls the expression of both the florescent protein mCherry, as well as a catalytically active form of TET1. In order to isolate cells that expressed a high level of active TET, as well as cells that had no active TET expression, florescence activated cell sorting was used.

To prepare cells for FACS, all cultures were washed twice, dissociated with Accutase, re-suspended in PBS, and pelleted as described in section *2.1.2 Cell culture, passaging, and harvesting*. Once pelleted, the supernatant was aspirated, and the cells were re-suspended in 500µL FACS buffer (PBS with 5% FBS and 1µM DAPI). Cell suspensions were mixed via pipetting to fully separate cells from one another, before being transferred to a new 15mL falcon tube through a 40µm strainer to remove any remaining cell clumps.

Cells were gated for five different parameters. Firstly, forward scatter height (FSC-H) vs forward scatter area (FSC-A) and side scatter height (SSC-H) vs side scatter area (SSC-A) were used to sort single cells and filter out any doublets. FSC-A vs SSC-A was used to select cells based on size, as any events outside these parameters are likely to be some kind of debris, either cellular or otherwise. SSC-A vs 4;6-diamino-2-phenylindole (DAPI) was used as a gate for live cells, as DAPI can only cross a compromised cell membrane to bind DNA (Kim et al., 2004), thus DAPI positive cells were excluded. Lastly, cells were sorted on mCherry expression. Cells were mCherry negative if fluorescence levels were equal to the negative control and were mCherry positive if fluorescence was 10^3 above baseline levels. Cells were

collected in 96 well plates containing 120 μL of GITC, resulting in cell lysis immediately after sorting.

2.7 Post-bisulfite adaptor tagging

Post-bisulfite adaptor tagging (PBAT) is a technique used to assess the methylation status across the genome, however, unlike whole genome bisulfite sequencing, PBAT can be performed with very low cell numbers (Miura et al., 2012). The PBAT technique can accomplish this by preforming bisulfite treatment before the addition of adaptor sequences for sequencing. This contrasts with other whole genome bisulfite sequencing methods which typically treat libraries with bisulfite after the addition of adaptors, which can lead to degradation of libraries.

Bisulfite conversion of genomic DNA (protocol described in 2.2.1 Bisulfite conversion), leaves it fragmented and single stranded. Thus, the first step was to synthesize a new strand, and attach the forward adaptor sequence. To do this, 19.5 μL of bisulfite treated DNA was mixed with 4 μL of first strand synthesis reaction mixture (Table 2.7). This mixture uses BioP5N7 as a primer for first strand synthesis. BioP5N7 contains a biotin molecule attached to the forward Otago index handle (Table 2.4), with a further seven random nucleotides. Samples were then incubated at 65°C for 5 minutes, then 4°C for 5 minutes, before addition of the DNA polymerase Klenow-exo (NEB, M0212L). All samples were then incubated at 4°C for 15 minutes, before the temperature was steadily increased at 6°C per minute until 37°C, where it remained for 90 minutes.

Table 2.7 PBAT first strand synthesis mixture

Reagent	**Volume**
dNTP (5mM) (BIO-39025)	1 μL
BioP5N7 (10 μM)	0.5 μL
10x NEB Buffer2 (B70025)	2.5 μL
Total	48 μL

After first strand synthesis the only remaining single stranded DNA is excess adapter. Thus, each sample was treated with 1μL of Exonuclease I (NEB, M0293S) to remove nucleotides from single stranded DNA and to prevent any further adapter dimer formation. The 37°C incubation was then continued for 60 minutes. Upon completion, 20 μL of PEG diluted solid phase reversible immobilization (SPRI) beads was added to each sample and mixed via pipetting. Samples were then left to incubate at room temperature for 1 minute to allow DNA to bind the beads before being placed onto a magnetic rack. The supernatant was aspirated, and samples were washed with 90% ethanol before being eluted into 100 μL of TE buffer.

Purified DNA fragments were then captured on streptavidin coated Dynabeads® (Thermo Fisher Scientific, 11205D). To do this, 10 μL of Dynabeads were first placed onto a magnetic rack and supernatant aspirated to remove the original solution. The beads were then removed from the magnet and re-suspended in 100 μL of binding and washing buffer (10mM Tris-HCl (pH7.5), 1mM EDTA, 2M NaCl). DNA fragments were then mixed with the streptavidin coated Dynabeads and incubated at room temperature for 20 minutes with rotation to allow for biotin capture. Samples were then placed on a magnetic rack, and supernatant removed and discarded.

To prepare samples for second strand synthesis and the attachment of the reverse adaptor, samples were washed twice with 100 μL of 0.1M NaOH to denature the first strand, before two more additional washes with 100 μL of 10mM Tris. Each sample was then re-suspended in 48 μL of second strand synthesis reaction mixture (Table 2.8). This mixture uses BioP7N7 as a primer for first strand synthesis. BioP7N7 contains a biotin molecule attached to the reverse Otago index handle (Table 2.4), with a further seven random nucleotides. Samples were then incubated at 95°C for 45 seconds before being immediately transferred to ice and the addition of 1 μL of DNA polymerase Klenow-exo (NEB, M0212L). Incubation was then continued at 4°C for five minutes, before rising to 37°C at 6°C per minute, where it remained for a further 90 minutes.

Table 2.8 PBAT second strand synthesis reaction mixture

Reagent	**Volume**
dNTP (5mM) (BIO-39025)	2 μL
BioP7N7 (10 μM)	2 μL
10x NEB Buffer2 (B70025)	5 μL
FSMQ water	39 μL
Total	48 μL

Upon completion, the supernatant was removed, and the beads were washed with 50 μL of TE buffer before being re-suspended in 12 μL of PBAT PCR mixture (Table 2.9). Amplification was carried out using the following thermal cycling parameters:

*98°C for 5min; 13*x (98°C for 80sec, 65°C 30sec, 72°C 30sec); 72°C for 5min; 8°C hold*

Once amplification was complete, each sample was run on a 1% agarose gel to assess sample quality, before being stored at -20°C until pooling in preparation for sequencing.

2 μL of each sample was pooled together into a single 0.2 mL PCR tube along with 2x the total volume of PEG diluted magnetic SPRI beads, only allowing fragment sizes above 200 base pairs to bind and eliminating all excess primer and polymerase. To ensure all DNA was bound to the beads samples were incubated at room temperature for 2 minutes before being placed onto a magnetic rack. The supernatant was aspirated, and beads were washed with 150 μL of 70% ethanol before being re-suspended in half the original starting volume. All samples were then sequenced in multiplex using the Illumina MiSeq system.

Table 2.9 PBAT PCR reaction mixture

Reagent	**Volume**
KAPA HiFi HotStart Uracil+ Mix (2x) (KK2801)	6 μL
Forward priming index (FPI) (10μM)	0.5 μL
Reverse priming index (RPI) (10μM)	0.5 μL
FSMQ water	5 μL
Total	12 μL

2.8 DNA sequencing

2.8.1 Sanger sequencing

Sanger sequencing was used to assess the single nucleotide polymorphism (SNP) profile in exon 2 and exon 5 of the *Impact* locus in the TET TKO mTET1-CD cell line. For amplicon sequencing, samples were first cleaned to remove excess dNTPs and enzymes. This was performed by mixing the PCR product with magnetic beads suspended in polyethylene glycol (PEG) at a ratio of 1:2 and incubating for 1 minute at room temperature to allow DNA to bind to the beads. Samples were then placed into a magnetic rack and the supernatant was aspirated and discarded. Each sample was washed with 70% ethanol twice, before being eluted into molecular grade water.

Once samples had been cleaned, 1 ng of DNA per 100 base pairs of template sequence was combined with 3.6 μM of primer in a 5 μL reaction. Sequencing was performed by the Genetic Analysis Service at the University of Otago and analysis was performed using Chromas software.

2.8.2 Illumina MiSeq

KCNQ1ot1 amplicons and PBAT samples were sequenced using the Illumina MiSeq System with the MiSeq Reagent Kit (MS-102-3001). This enables high throughput sequencing of multiple samples at once, speeding up the sequencing process tremendously.

The MiSeq Reagent Kit was stored at -20°C. In preparation for sequencing the reagent cartridge was removed from storage at least 1 hour before sequencing and

placed in room temperature FSMQ water to thaw. The hybridization buffer (HT1) was thawed and stored at 4°C until use.

To prepare samples for sequencing cleaned libraries (using PEG diluted SPRI beads for PBAT libraries or by gel extraction for amplicon libraries) were diluted to 4nM with FSMQ water. DNA was then denatured by combing 5µL library with 5µL of freshly prepared 0.2nM NaOH and mixing by vertex for 1-minute, before a 5-minute incubation at room temperature. Libraries were then diluted by adding 990µL of HT1 resulting in a 20pM denatured library. The 20pM library can then be diluted further to obtain a specific cluster density. To obtain a cluster density of 800k/mm^2, 240µL of 20pM library was mixed with 360µL of HT1 to get a final concentration of 8pM.

Finally, 600µL of sample library was loaded into the reagent cartridge reservoir. The reservoir, flow cell, and incorporation buffer were then loaded into the MiSeq System for sequencing to begin.

2.9 Bioinformatics

In order to process the large amount of data generated by high-throughput multiplex sequencing a bioinformatics pipeline was developed. This involved quality checking raw reads and trimming the adapter sequences, separation of maternal and paternal alleles, mapping reads to a reference sequence, and generating easy to understand graphs from data-heavy samples (Figure 2.3).

2.9.1 Quality control and adapter trimming

2.9.1.1 FASTQC

FASTQC is a quality control software developed by the bioinformatics group at Babraham Institute, England. This software provides an overview of raw data quality, including per base and per sequence quality scores, sequence length distributions, adapter content, and the presence of overrepresented sequences. Upon receiving sequencing results, all samples were checked in FASTQC for an indication of data quality and to identify any areas where problems might exist. This tool would be continually used to examine read quality throughout data analysis to monitor the quality of samples. For example, samples with high adapter content would be re-examined after adapter trimming to ensure all adapter sequences had been removed correctly.

2.9.1.2 Adapter trimming

Adapter sequences are required for cluster formation during sequencing, however, they must be removed prior to analysis. This is because unlike the template DNA, they have not undergone bisulfite conversion. Thus, in order for bisulfite conversion efficiency calculations to be accurate these sequences must be removed. To achieve this the 'TrimGalore' program was used in conjunction with the UNIX command line. In essence, this program could loop through all the reads of a given sample and trim off a predetermined number of nucleotide calls equal to the length of the adapter sequence.

2.9.2 Parent specific allele isolation and BISMARK mapping

2.9.2.1 Parent specific allele isolation

In order to (a) determine the number of maternal and paternal alleles per sample, and (b) assess parent specific methylation patterns, reads pertaining to either parent were

isolated and assessed individually where applicable (i.e TET TKO samples had no parent identifying features and thus separation was not possible).

To do this, regular expression patterns were identified and isolated using the grep command in the terminal command line. To preserve the FASTQ format of the files, one line above the regular expression, and two lines below, were also extracted. The following command was utilized to achieve this:

```
grep -A 2 -B 1 '[regular expression]' [INPUT.fq] | sed '/^--$/d' > [output.fq]
```

Parent-of-origin allelic makers are often situated in the middle of the amplicon; thus, their isolation has the added benefit of removing short, low quality reads that do not contain these markers but were not removed during the adapter trimming phase.

2.9.2.2 BISMARK Mapping

Once reads had been trimmed and parent specific reads had been isolated, the BISMARK mapping program was used to map reads to a reference sequence, allowing for identification of specific CpG methylation profiles. Output files were then extracted and processed by the SAMtools, converting them into a binary arrangement where each CpG location within a read was marked as either methylated or unmethylated.

2.9.3 Graph production

A custom R-script was developed to create easily readable images of the data. In short, this script would assess the methylation status at each CpG for each read, calculate the overall mean methylation across all reads in that sample, and produce a ‘salt and pepper’ heat map where methylated CpGs are displayed as a black square and unmethylated CpGs are displayed as a white square.

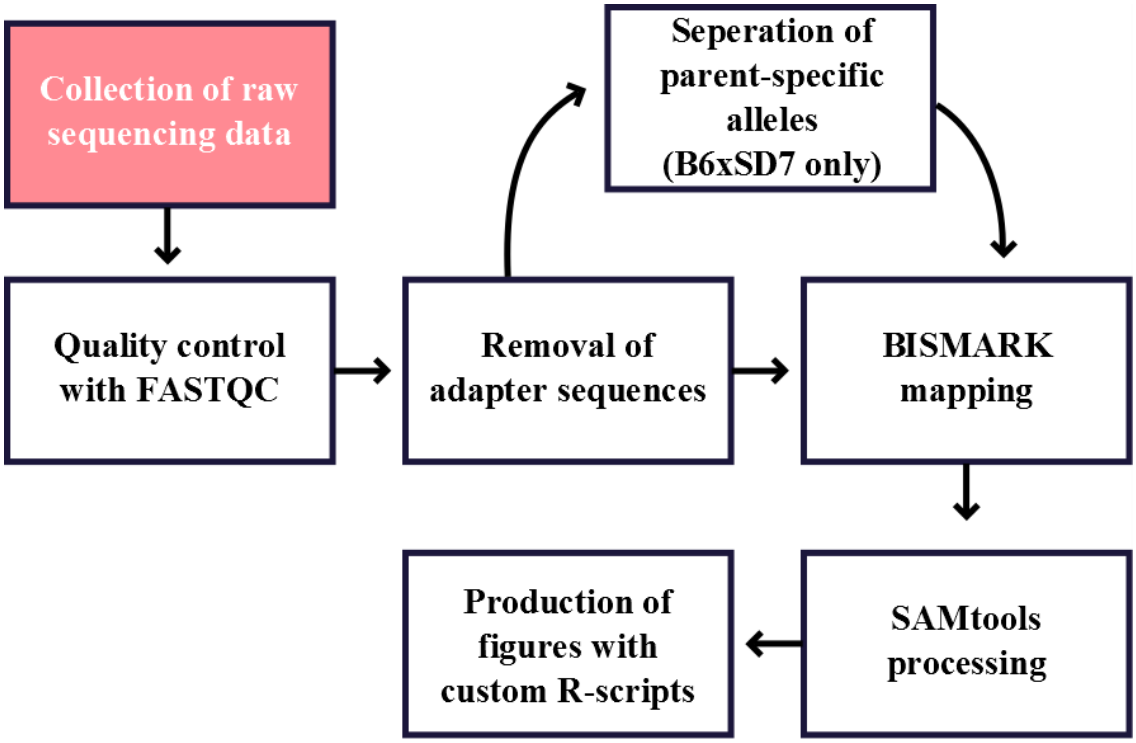


Figure 2.3 Bioinformatics Pipeline

Schematic showing the pipeline developed for analysis of bisulfite amplicon sequencing from raw sequencing data to the production of figures.

3 CHAPTER THREE: RESULTS

Inducing global DNA demethylation by enhancing TET enzyme activity is a useful technique to improve the reprograming efficiency of differentiated cells to a naïve pluripotent state (Hore et al., 2016). Nevertheless, too much TET activity could cause cells to lose methylation at imprinting control regions, resulting in loss of imprinting and biallelic gene expression.

To test this possibility, mouse embryonic stem cells were cultured under 2i conditions and treated with varying concentrations of two antioxidants: ascorbate and hydroquinone. Following optimisation of PCR amplification for two imprinted loci, high-throughput bisulfite amplicon sequencing was used to assess the methylation patterns at the KCNQ1ot1 imprinted locus. While technical difficulties such as the detection of clonal amplification, particularly in the TET triple knockout cell line, hindered this study, I was able to show that the antioxidant ascorbate causes a dose dependant loss of methylation. However, most interestingly, at the higher concentration of 100 ng/μL of ascorbate, no loss of methylation was observed.

3.1 PCR optimization

To efficiently quantify methylation in a large number of samples, I performed multiplex bisulfite amplicon sequencing using the Illumina MiSeq platform. The power of multiplex library preparation lies in the paired-end 8-bp indexes that flank the region of interest – by using a unique combination of indexes for each sample, I was able to analyse >100 samples simultaneously. The indexes, as well as Illumina adapters (for

binding the sequencing flow-cell) were added to template DNA during a 2-step PCR amplification process (Figure 2.2). Due to the extreme length of the template several PCR optimisation experiments had to be performed in order to consistently generate high quality amplicon libraries.

3.1.1 *Cell culture*

Both V6.5 and B6xSD7 cell lines were chosen as candidate cell lines due to the parent-of-origin specific allelic differences at the *Impact* and KCNQ1ot1 loci, respectively (Figure 2.1). The V6.5 cell line was observed to regularly form compact circular colonies typical of naïve embryonic stem cells (Ying et al., 2008) when cultured in either the FBS-based serum + 2i/LIF media or the chemically defined N2B27 + 2i/LIF media. There was no apparent change in in cell growth between either media condition. This is in contrast to the B6xSD7 cell line which was extremely sensitive to the N2B27 + 2i/LIF media, which often crashed within a few days of seeding. The chemically defined nature of N2B27 + 2i/LIF media offers greater control over media composition. Thus, the V6.5 cell line, and therefore the *Impact* locus, was the ideal first choice.

3.1.2 *Impact* PCR optimisation

The *Impact* locus has two SNPs within the protein coding region, one in exon 2 and another in exon 5 (Figure 2.1b). To assess the imprinting stability at this locus, I attempted to examine the relative expression patterns of each allele using reverse transcriptase PCR (RT-PCR). In order to determine the optimal annealing temperature, I performed a gradient PCR experiment using four different annealing temperature between 54°C and 65°C. I used tagged primers for *Impact* exon 2 and exon 5 cDNA amplification (Table 2.2) with the Phusion® High-Fidelity DNA Polymerase (Table

2.6), and cDNA from previously extracted V6.5 TNA. The following thermal cycling parameters were used:

98°C, 2 min; 34(98°C, 10 sec; 54-65°C, 10 sec; 72°C, 20 sec); 8°C ∞*

The correct band size for both exon 2 (272 base pairs) and exon 5 (291 base pairs) confirmed successful amplifcation across all anneling tempretures. In addtion, non-specific amplifcation was not seen within any of the anneling temperatures (Figure 3.1), indicating that any of the tested temperatures are adequate for amplicon library generation.

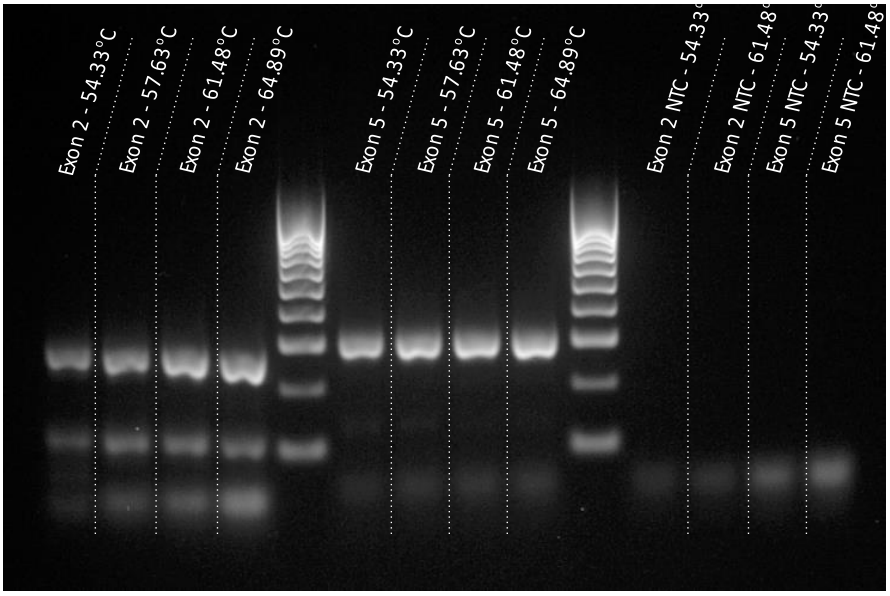


Figure 3.1 *Impact* gradient PCR for round one amplification optimization

Agarose gel electrophoresis analysis of the *Impact* exon 2 and exon 5 gradient PCR. Successful amplification for each region of interest across all temperature ranges is confirmed by the correct band size for both exon 2 (272 base pairs) and exon 5 (291 base pairs). Additional bands at 110 base pairs are presumed to be primer dimer.

Once sucessful amplification had been observed with sequence specific primers, I attempted to attach unique indexes. It was possible that the indexing primers would have a different optimal annealing temperature than the sequence specific primers. Therefore I performed another gradient PCR experiment using the previous conditions, along with the addition of a second amplifcation cycle using unique indexing primers (supplementary Table 7 and 8). The following thermal cycling paramters, with the bolded 4°C pause step indicating the addtion of the unique indexing primers:

98°C, 2 min; 28(98°C, 10 sec;* 65°C*, 10 sec; 72°C, 30 sec);* ***4°C*** *∞****;*** *5*(98°C, 10 sec; 54-*65°C*, 10 sec; 72°C, 30 sec); 8 ∞.*

The addition of an extra band 65 base pairs above the original fragment height creates a 'double band' on the gel. This indicates the successful incorporation of indexes into most, but not all of the amplicons and was used as a marker for correct library generation. Upon completion of the indexing PCR only single bands were present, suggesting that the indexes had not been successfully incorporated at any of the temperature ranges (Figure 3.2a). It was hypothesized that the low number of cycles were responsible for this, thus the number of cycles was increased to 8. This resulted in apparent successful incorporation of indexes into exon 5 sequences at 61.48°C, but interestingly, not at the lower annealing temperature of 57.63°C. No amplification was observed for exon 2 (Figure 3.2c). Regardless, the *Impact* exon 5 libraries were sequenced. Unfortunately, for an unknown reason, these libraries did not return any reads. Therefore, due to time constraints and the availability of the B6xSD7 cell line the decision was made not to continue with the V6.5 cell line and *Impact.*

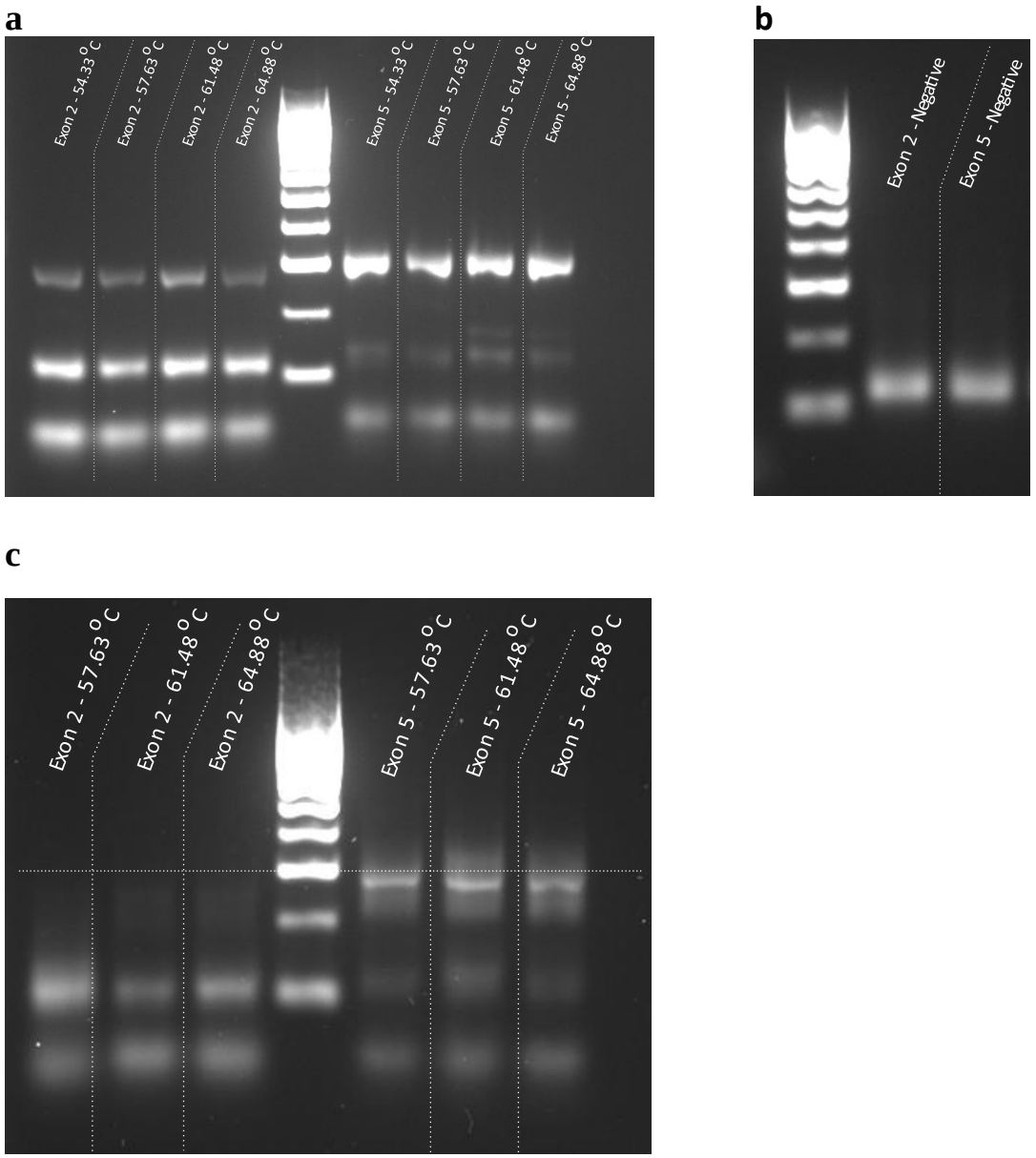


Figure 3.2 *Impact* gradient PCR for round two amplification optimization

(a) Agarose gel eletorpheisis anaylsis for indexing gradeint PCR with 5 cycles post the additon of indexing primers. The presence of only one band (272 base pairs for exon 2, 291 base pairs for exon 5) indicates the unsuccessful addtion of unique indexes to the amplicons. **(b)** Negative control for exon 2 and exon 5 indexing gradient PCR. The band at 110 base pairs is presumed to be primer dimer. **(c)** Agarose gel electrophoresis anaylsis for indexing gradient PCR with 8 cycles post the addition of indexing primers. Amplification of the correct band size of 272 base pairs was not seen in exon 2. Exon 5 contains two bands at 291 base pairs and 256 base pairs indicates the presence of the orginal amplicon and amplicon + index respectively.

3.1.3 KCNQ1ot1 PCR optimisation

The KCNQ1ot1 primers (Table 2.1) have previously been employed in the Hore laboratory. Successful amplification and incorporation of indexes has been achieved using the KAPA HF DNA polymerase mix (Table 2.5) with the following thermal cycling parameters.

92C, 2min; 25(92C, 10s; 59C, 10s; 72C, 30s); 72C, 5min;* ***4C, infinite****; 7*(92C, 10s; 59C, 10s; 72C, 30s); 72C, 5min; 4C, infinite.*

Figure 3.3a shows successful amplification of the KCNQ1ot1 locus using bisulfite treated B6xSD7 DNA. The presence of two bands, although faint, indicates successful incorporation of the unique indexes to the template sequence. Sequencing with the Illumina MiSeq system was used to confirm the successful index incorporation. However, this protocol proved inconsistent, especially when using low concentration DNA as a template. Increasing the number of cycles in the second round of PCR to ten from seven improved the constancy and reliability of this protocol (Figure 3.3b).

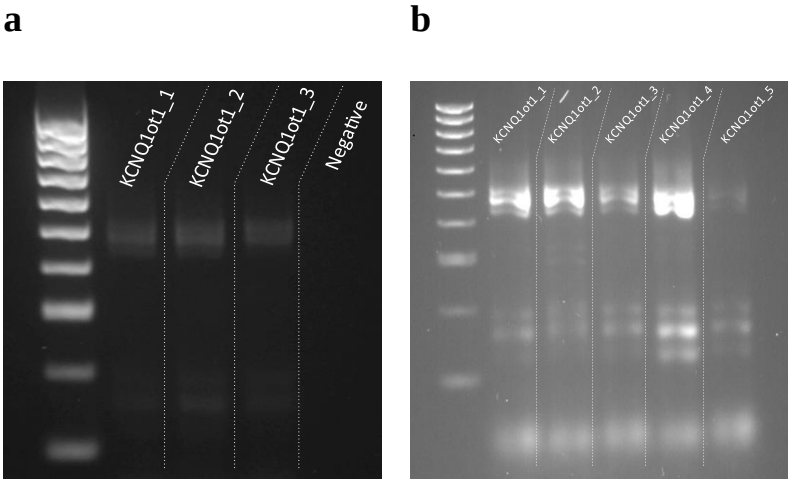


Figure 3.3 Incorporation of unique indexes into KCNQ1ot1 amplicon

Agarose gel electrophoresis analysis of the KCNQ1ot1 amplicon (446 base pairs) after two successive rounds of PCR, where **(a)** Seven additional cycles were used to incorporate indexes, n=3, or **(b)** Ten additional cycles were used to incorporate indexes, n=5

3.2 B6xSD7

3.2.1 B6xSD7 Cell line establishment and antioxidant treatment

The B6xSD7 cell line was first grown in several different culture media. These included:

- Serum + 2i/LIF
- Serum + LIF (in the absence of 2i)
- N2B27 + 2i/LIF

In serum + LIF conditions, the cells formed flattened, spread out colonies. However, in Serum +2i/LIF and N2B27 +2i/LIF conditions well defined circular colonies typical of naïve embryonic stem cells were observed (Ying et al., 2008). The B6xSD7 cell line appeared to be sensitive to N2B27 + 2i/LIF, as cultures would often arrest and subsequently crash after a few days in this condition. However, there appeared to be no substantial difference in growth rate between serum + LIF and serum + 2i/LIF conditions. Considering this, serum + 2i/LIF was used for antioxidant treatments because it induces naïve pluripotency, as opposed to the primed pluripotency of classical ES cells grown in the absence of 2i (Ying et al., 2008, Nichols and Smith, 2009). This is of particular importance as classically grown ES cells have varied gene expression patterns within the population, while naïve cells have homogenous gene expression (Wray et al., 2010).

B6xSD7 cells were incubated in serum + 2i/LIF conditions for 168 hours (7 days) before antioxidant treatment to induce naïve pluripotency and subsequent homogenous gene expression. ES cells were treated with various concentrations of ascorbate and hydroquinone (Table 3.1) for 72 hours. Hore et al, (2016) demonstrated that either 50 ng/μL of ascorbate or 50μM of hydroquinone

supplementation is sufficient to enhance TET activity *in vitro*. Thus, these concentrations served as the midpoint for treatment range. However, concentrations above 50 μM of hydroquinone proved toxic and no cell growth was detected. Therefore, 50 μM was set as the top concentration, and the range was scaled accordingly. Ascorbate treated samples showed inhibited cell growth at concentrations over 100 ng/μL, with no live cells being detected in 400 ng/μL of ascorbate.

Table 3.1 B6xSD7 antioxidant treatment range

B6xSD7 antioxidant treatment range								
Ascorbate (ng/μL)	0	6.25	12.5	25	50	100	200	400
Hydroquinone (μM)	0	0.78	1.56	3.13	6.25	12.5	25	50

3.2.2 B6xSD7 - Bisulfite amplicon sequencing of KCNQ1

I used bisulfite amplicon sequencing to assess the methylation patterns at the KCNQ1ot1 imprinted loci after treatment of each antioxidant. Two successive rounds of PCR, as described in sections 2.4 and 3.1.2, were employed to amplify and attach unique indexes to the KCNQ1ot1 region at each antioxidant concentration (Table 3.1). Three independent biological replicates were used. Amplification was not observed at all concentrations – specifically, only products of incorrect size were found when media was supplemented with 200-400 ng/μL of ascorbate (Figure 3.4). Samples treated with hydroquinone had variable levels of amplification, however bands at the correct amplicon size were detected at each concentration level.

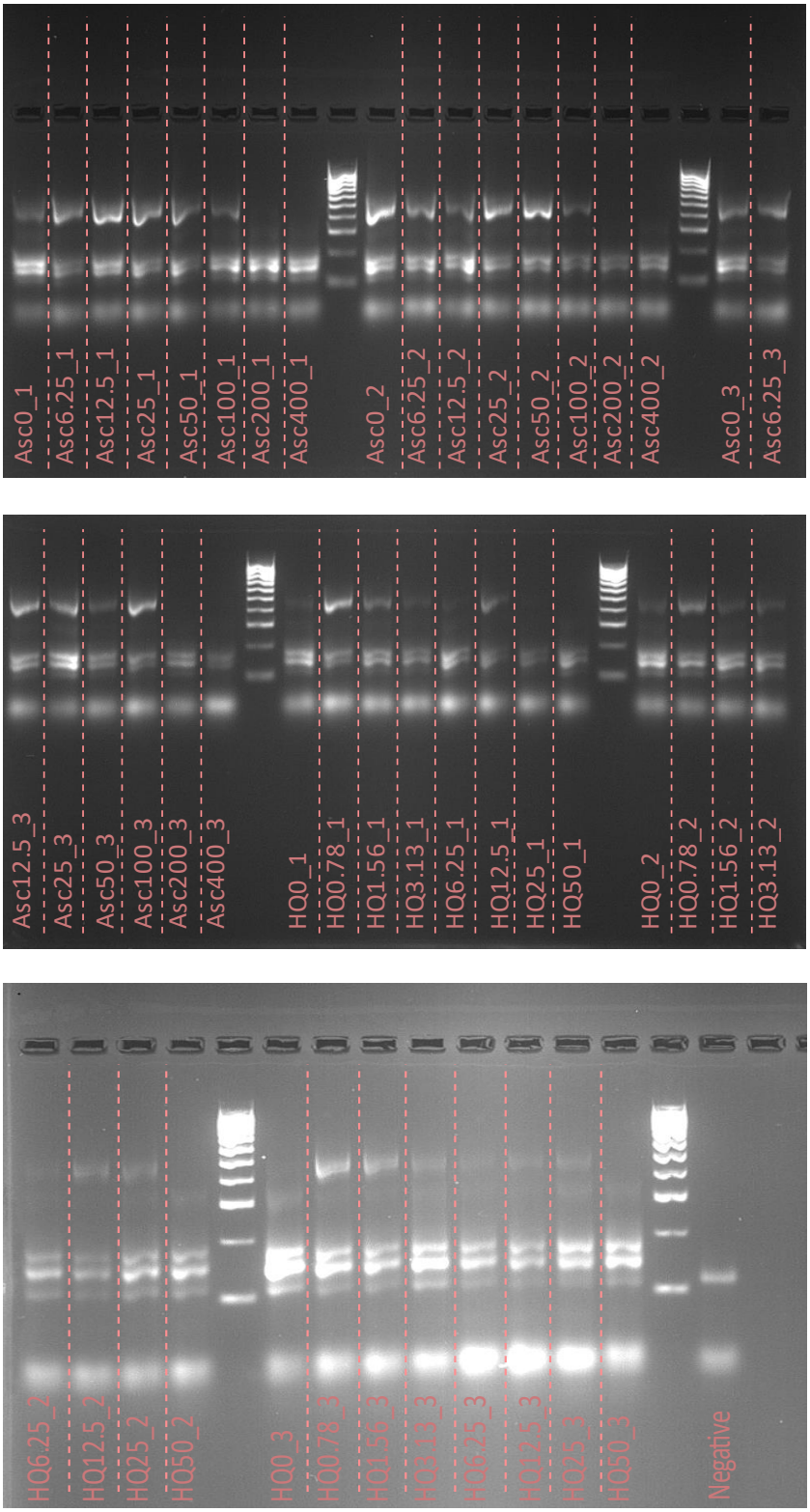


Figure 3.4 B6xSD7 ascorbate and hydroquinone amplicon library generation

Agarose gel electrophoresis analysis of amplicon libraries after cell treatment with either ascorbate or hydroquinone, n=3. 'Double bands' at 446 base pairs indicate KCNQ1ot1 amplicons with and without indexes attached. The three bands between 100 and 200 base pairs are indicative of primer dimer.

Amplicons were sequenced in multiplex with the Illumina MiSeq system. The number of reads per samples for ascorbate treated cells per was extremely variable within biological replicates (Figure 3.5). This is expected however, as the number of reads parallels the respective band intensities seen in Figure 3.4, suggesting that the number of reads is proportional to the amount of DNA being sequenced.

Unsurprisingly, less than 50 reads were recorded for any of the replicates treated with either 200 ng/µL or 400 ng/µL of ascorbate. However, most importantly, all other samples had over 500 reads, and all but four had more than 1000 reads.

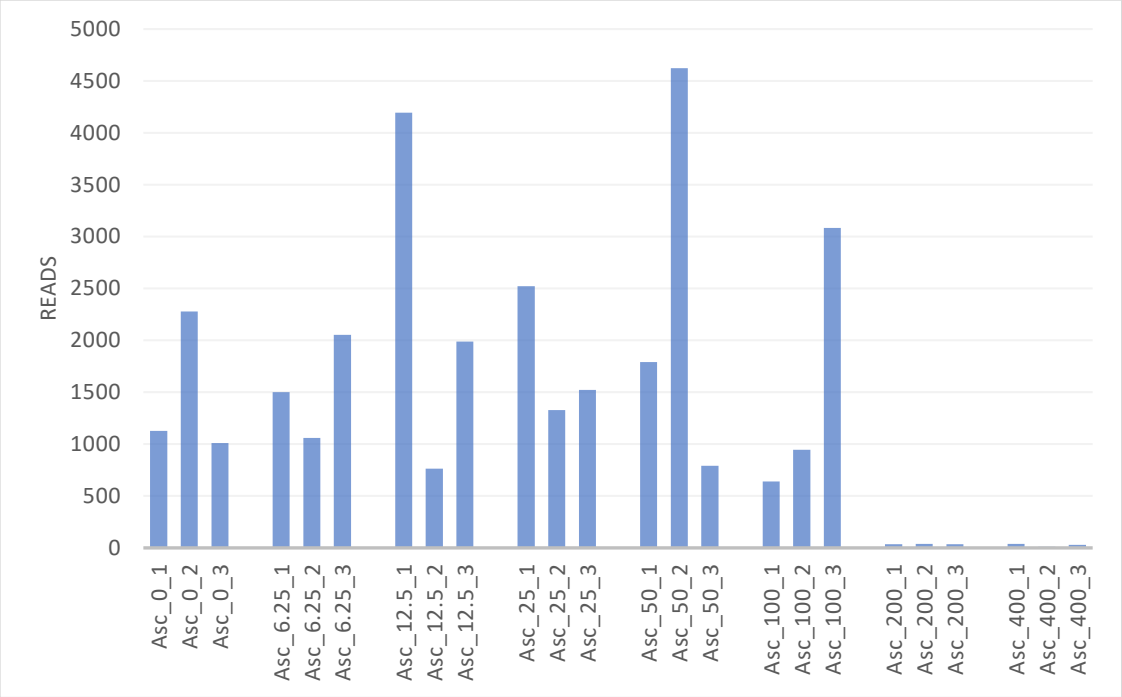


Figure 3.5 Read count from ascorbate treated cells.

Bar chart displaying the read count for all samples treated with ascorbate (Asc), where the first number indicates the concentration of ascorbate (ng/µL) and the second is the number of the replicate.

All hydroquinone treated samples, with the exception of those treated with 0.78 µM hydroquinone, produced very low read counts (Figure 3.6). Interestingly, the HQ0 (0µM hydroquinone) replicates and Asc0 replicates (0ng/µL ascorbate) had drastically different read counts despite having the same culture conditions. This suggests that the DNA extraction and bisulfite conversion, which occurred separately between the two treatment samples, resulted in decreased DNA yield in the hydroquinone samples, but not the ascorbate samples.

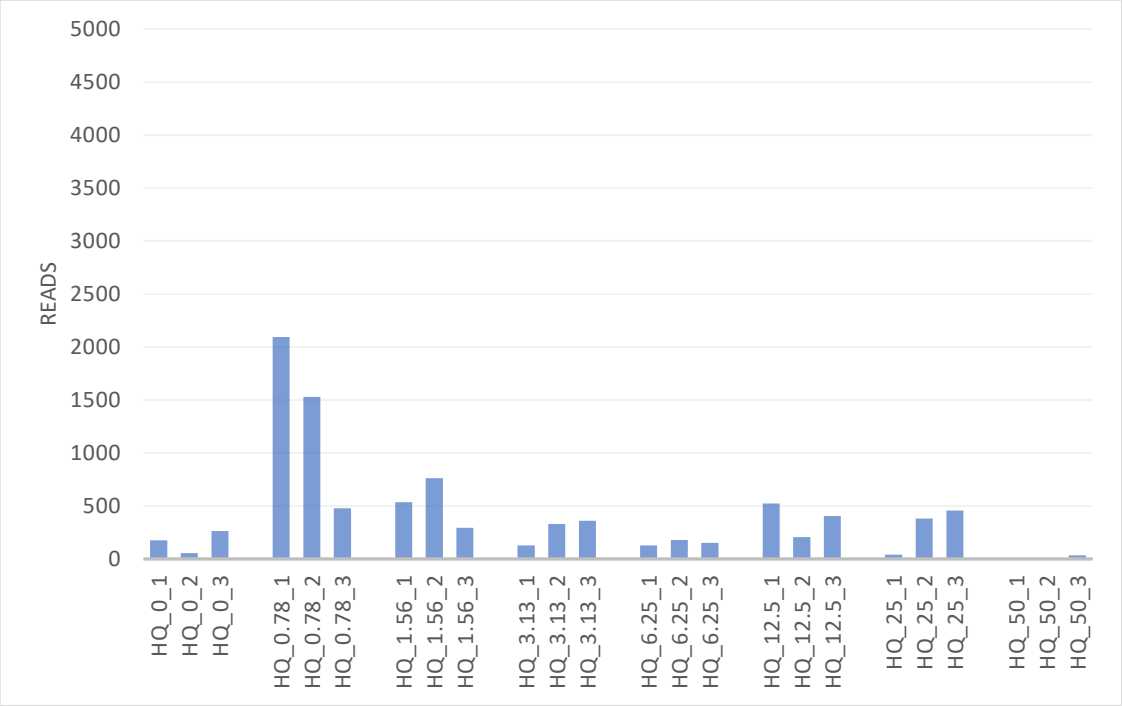


Figure 3.6 Read count from hydroquinone treated cells

Bar chart displaying the read count for all samples treated with hydroquinone (HQ), where the first number indicates the concentration of ascorbate (µM) and the second is the number of the replicate.

3.2.2.1 Bisulfite conversion efficiency

The bisulfite chemical preferentially deaminates unmodified cytosines to uracil. Conversion efficiency can be altered by reaction conditions and reagent quality; thus, it is important to assess conversion efficiency for each reaction. The majority of DNA methylation in vertebrates is found within the CpG context, thus non-CpG sites can be used to approximate conversion efficiency provided DNA methylation levels are low in this context. Non-CpG methylation is enriched within classically grown mammalian ES cells, due to high DNMT3a and DNMT3b enzyme expression (Ramsahoye et al., 2000, Lee et al., 2017). However, downregulation of all DNMT enzymes contributes to low methylation levels in naïve ES cells grown under 2i conditions (Choi et al., 2017, von Meyenn et al., 2016, Lee et al., 2014b, Habibi et al., 2013, Ficz et al., 2013), resulting in non-CpG methylation levels of ~0.1%. Thus, non-CpG conversion rate is likely a good measure of overall bisulfite conversion efficiency for naïve ES cells.

Both hydroquinone and ascorbate treated samples achieved extremely high conversion efficiencies (Figure 3.7), with all but one sample (HQ_50) showing conversion rates of 99.7% or higher. Considering the 0.1% non-CpG methylation reported by Habibi et al, (2014), it appears there was almost complete bisulfite conversion across all samples, suggesting that the CpG methylation patterns observed in these samples are an accurate representation of the true CpG methylation patterns.

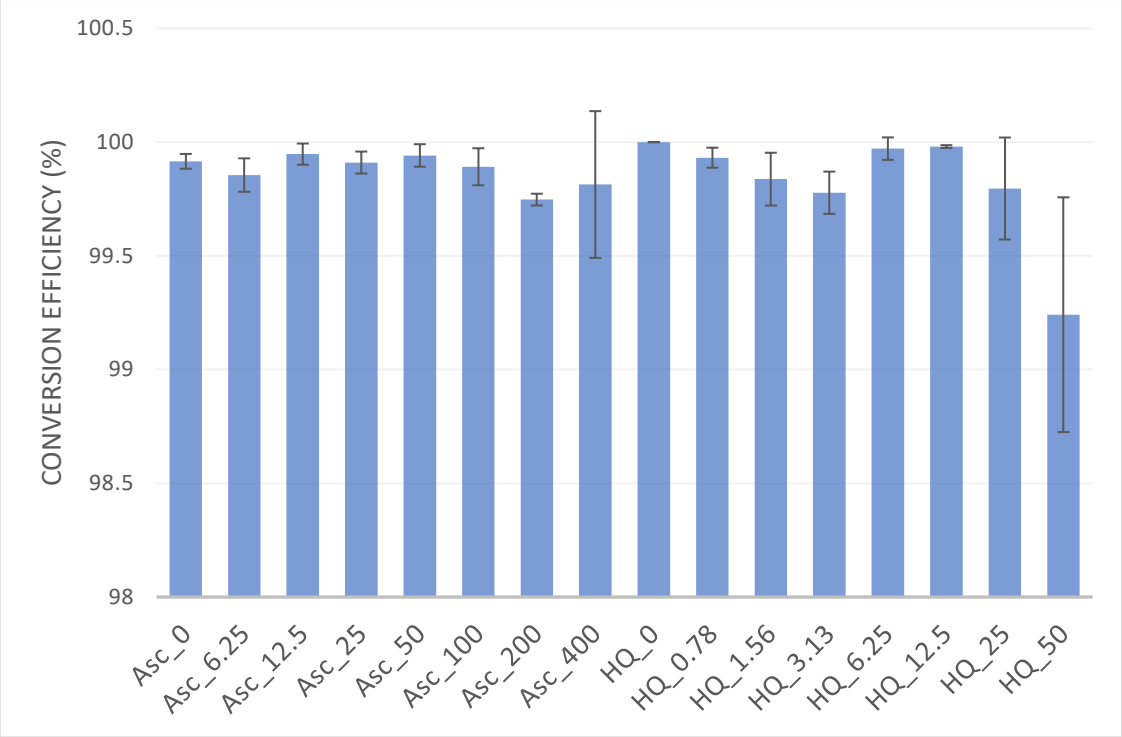


Figure 3.7 B6xSD7 Bisulfite conversion efficiency

The mean bisulfite conversion efficiency for ascorbate (Asc) and hydroquinone (HQ) treated samples, as determined by the ratio of unconverted to converted cytosines in the non-CpG context (n=3, ±S.D.). The final concentration of supplemented Asc or HQ (ng/mL or µM respectively) is indicated for each sample.

3.2.2.2 Allele Frequencies.

The relative frequencies between maternal (B6) and paternal (SD7) alleles in the B6xSD7 samples were recorded to (a) determine allele specific methylation patterns, and to (b) determine if any amplification bias had occurred. To achieve this, the allelic differences (Figure 2.1) were used to isolate parent specific reads. Interestingly, some samples (Asc0, Asc6.25, and Asc25 + HQ0.78 + HQ6.25) had significantly more paternal reads than maternal reads, while the remaining samples had no significant difference. Ascorbate treated samples tended to trend towards more paternal alleles than maternal, with the exception of Asc200 and Asc400; which also possess a high

amount of variation. It is possible that the difference seen in the latter two samples is due to (a) very low read counts, and (b) 'clonal expansion' during PCR amplification (Figure 3.9). Clonal expansion is where one parental allele gets preferentially amplified because of low template concentration. Clonal amplification can also be detected by analysing the diversity of methylation patterns within a population of fragments which is discussed later in the results (section 3.2.3).

Hydroquinone treated samples also appeared to trend towards higher paternal allele reads than maternal reads. While only two were statistically different from expected (HQ0.78 and HQ6.25), most hydroquinone samples showed high variation between replicates indicating skewed amplification indicative of clonal expansion (Figure 3.8b). The paternal allele is six base pairs shorter than the maternal allele. This difference may account for the somewhat apparent skew towards the paternal allele.

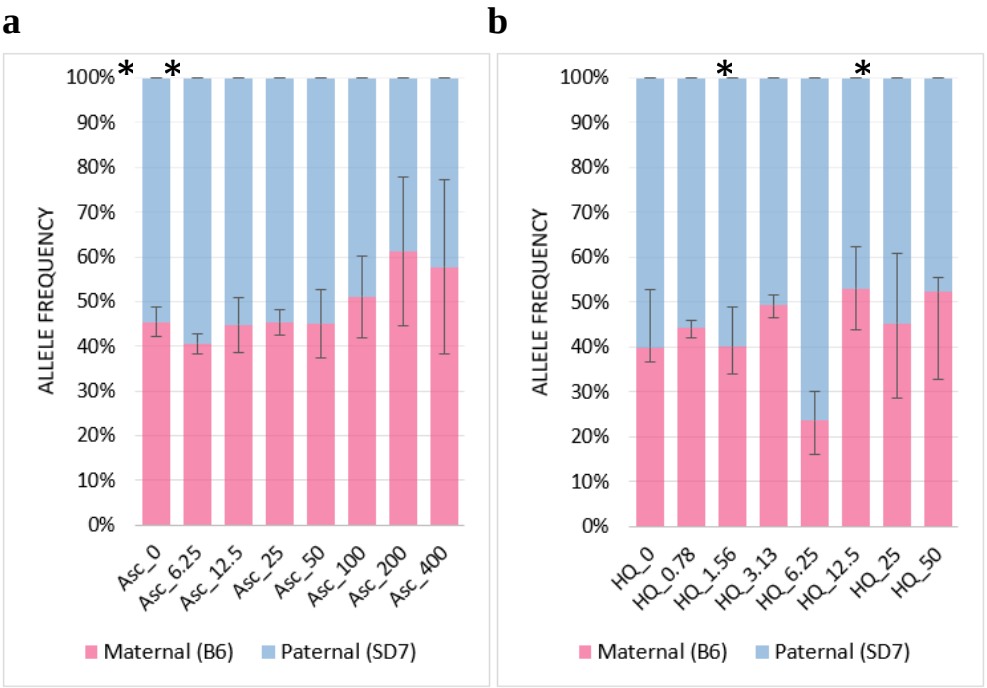


Figure 3.8 KCNQ1ot1 allele frequencies in B6xSD7

Relative parental allele frequencies after sequencing in the B6xSD7 cell line. The maternal (pink) and paternal (blue) alleles for **(a)** ascorbate and **(b)** hydroquinone treated samples is shown (n=3, ±S.D.). Two-way students t-test.

3.2.3 Clonal expansion

Amplicon bisulfite sequencing is prone to clonal amplification if low amounts of starting material are used. Given skewed amplification of parental alleles from B6xSD7 template, I suspected that clonal amplification may be prevalent in my dataset, particularly for the high ascorbate concentrations (where there was a lot of cell death) and the hydroquinone samples. To explore this further, I examined the patterns of methylation for each sample and identified where clonal amplification was likely occurring, as can be seen in Figure 3.9.

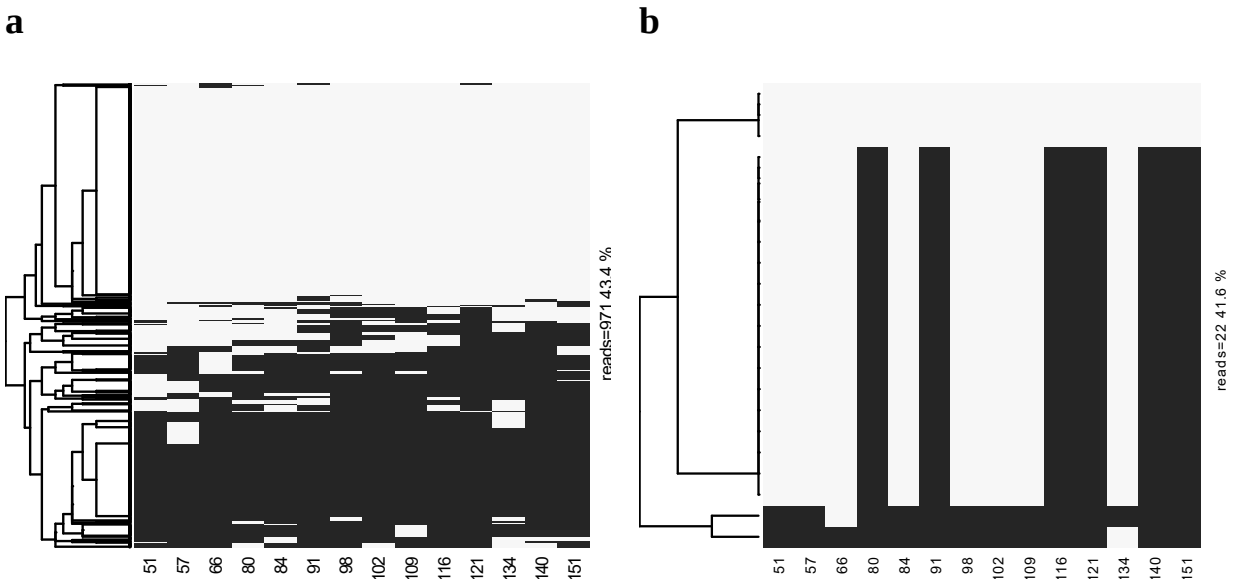


Figure 3.9 Clonal Expansion

Clonal expansion of the maternal allele of B6xSD7 after ascorbate treatment. **(a)** An example of a library with no clonal expansion after amplification (25 ng/µL of ascorbate), **(b)** An example of a library with clonal expansion after amplification (400 ng/µL of ascorbate). Clonal amplification is exemplified by repeated patterns throughout the entire library.

3.2.4 Effects of ascorbate on KCNQ1ot1

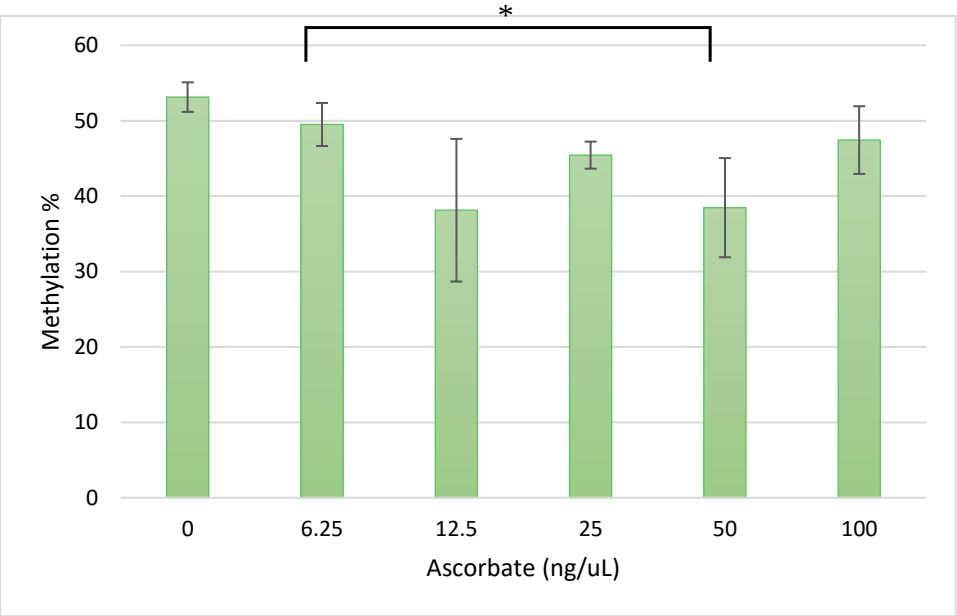
In this experiment, the methylation patterns at the KCNQ1ot1 locus were examined after ascorbate treatments across several concentration ranges (Table 3.1). KCNQ1ot1 is a maternally controlled region, meaning the maternal allele is methylated, while the paternal is unmethylated (Royo and Cavaille, 2008).

By isolating the reads from the maternal alleles only, I compared the methylation patterns at each level of ascorbate treatment. Asc200 and Asc400 both showed signs of major clonal amplification, as seen by consistently repeated methylation patterns and as a result were excluded from any further analysis. Surprisingly, even with no ascorbate treatment, there was only a maximum of 55% CpG methylation across all samples and all three biological replicates. 2i conditions are known to induce demethylation at ICRs, however, this usually takes many passages and prolonged culture (Choi et al., 2017), as opposed to the 10 days of culture in 2i used in these experiments, underlining how unpredictable imprint stability can be under 2i conditions. .

Interestingly, ascorbate treatment appeared to cause a decrease in DNA methylation, although the effects were subtle. All levels of ascorbate supplementation except 100 ng/nL caused significant demethylation when compared to the control individually, however, significance was lost for all samples except the 25 ng/mL level following Bonferoni correction for multiple comparison (Figure 3.10). One unexpected result was that a clear response to ascorbate was not obvious. I hypothesized that higher concentrations of ascorbate would result in more rapid demethylation, however ES cells with 100 ng/µL ascorbate had the second highest amount of methylation behind 6.25 ng/µL of ascorbate when compared to the control, despite having twice the amount of

ascorbate as the next highest treatment. This data suggests that there may be an optimal range somewhere between 12.5ng/μL - 100 ng/μL where demethylation can occur under ascorbate supplementation. Indeed, Blaschke et al, (2013) report no significant demethylation after ascorbate treatment at 100 ng/μL.

a



b

Ascorbate Treatment	P-Value	Significant before Bonferroni correction (p<0.05)	Significant after Bonferroni correction (p<0.01)
0 ng/µL	1	-	-
6.25 ng/µL	0.1435	×	×
12.5 ng/µL	0.0546	✓	×
25 ng/µL	0.0074	✓	✓
50 ng/µL	0.0208	✓	×
100 ng/µL	0.1139	×	×

Figure 3.10 Effect of ascorbate treatment on KCNQ1ot1 methylation

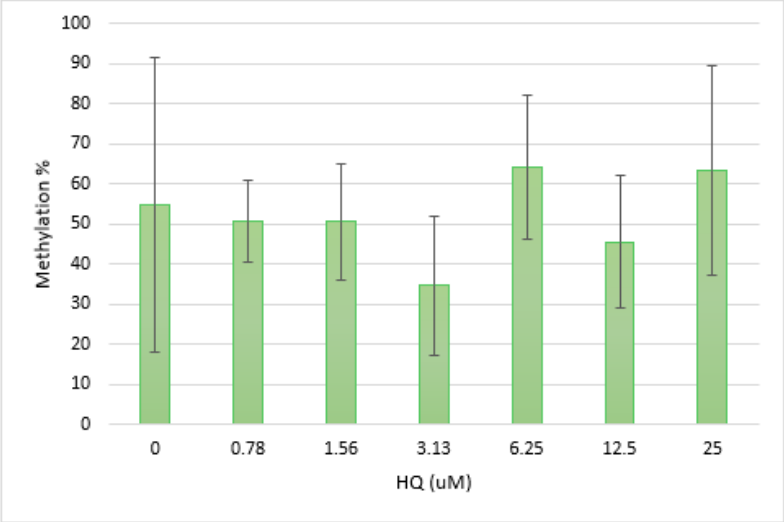
(a) Bar graph depicting overall methylation percentage at KCNQ1ot1 (n=3, ±S.D.) * adjusted p-value < 0.01. **(b)** All treatments were compared to 0 ng/µL using a two-tailed unpaired t-test with Bonferroni correction for multiple comparisons.

3.2.5 Effects of hydroquinone on KCNQ1ot1

The effect of hydroquinone on maternal methylation stability at the KCNQ1ot1 locus was also tested. Similarly, to the ascorbate treated cells, loss of methylation was present in maternal alleles, even in the control, suggesting prior loss of methylation.

Unfortunately, all samples presented with signs of clonal amplification during sequencing, as shown by repeated methylation patterns and low read diversity (Supplementary Figure 4), as well as high variation in parental allele frequency (Figure 3.8). Unsurprisingly, due to the extremely high variation between biological replicates (Figure 3.11a) no statistically significant effect was observed as a result of hydroquinone treatment. By the same token, this data cannot be used to infer the opposite conclusion; I cannot be sure that hydroquinone does not have an effect on ICR stability due to the low quality of the data.

a **b**



Hydroquinone (µM)	p-value
0	-
0.78	0.8647
1.56	0.8626
3.13	0.4383
6.25	0.7131
12.5	0.7123
25	0.7536

Figure 3.11 Effect of hydroquinone treatment on KCNQ1ot1 methylation

(a) Bar graph depicting overall methylation percentage at KCNQ1ot1 (n=3 ± S.D.). High variance is an additional indicator of clonal expansion. **(d)** All treatments were compared to 0 ng/µL using a two-tailed unpaired t-test with Bonferroni correction for multiple comparisons. * adjusted p-value < 0.01.

3.3 TET triple Knockout

3.3.1 TET TKO mTET1-CD cell line establishment

Up until this point it had been assumed that antioxidant treatment was eliciting changes by enhancing TET activity. In order to explore this further, the TET TKO mTET1-CD cell line was simultaneously tested. This cell line does not possess any endogenous TET protein, as all three TET variants contain loss of function mutations in the catalytic domain (Hu et al., 2014). However, in addition to this, the TET TKO mTET1-CD line has been transfected with functional TET1, as well as the reporter protein mCherry, under the control of a doxycycline inducible promotor; allowing for rescued TET expression in a controlled manner (Hore et al., *unpublished*).

Similarly to the B6xSD7 line, the TET TKO mTET1-CD cells were initially grown in three different media:

- Serum + 2i/LIF
- Serum + LIF (in the absence of 2i)
- N2B27 + 2i/LIF

There were no observable differences in growth between serum + 2i/LIF or serum + LIF in the absence of 2i. However, the growth rate was inhibited by N2B27 + 2i/LIF, but no culture crashes were observed like in B6xSD7. As a result, both serum + 2i/LIF and N2B27 + 2i/LIF media were used in the current experiment.

Antioxidant treatment in the TET TKO line covered a similar range as B6xSD7, but fewer increments were used (Table 3.2). In addition, 1 μM of doxycycline was added to the culture media to induce TET1 expression.

Table 3.2 TET TKO mTET1-CD antioxidant treatment range

TET TKO mTET1-CD antioxidant treatment range			
Ascorbate (ng/μL)	0	12.5	50
Hydroquinone (μM)	0	6.25	25

3.3.2 Parental origin identification of the Impact locus in the TET TKO cell line

Prior to the beginning of this project, it was unknown if the TET TKO cell line contained any allele specific differences at any ICRs. However, it was possible there were SNPs at the *impact* locus, similar to those in the V6.5 cell line (Figure 2.1). This is because the TET TKO line has a mixed genetic background, primarily of the 129/sv strain, but also containing some C57BL6 (B6) (Figure 3.12). It was possible that the *Impact* locus contains one B6 allele and one 129/sv allele, therefore providing me a way to distinguish the maternal and paternal alleles apart.

To determine if this was the case, the genomic region of exon 2 and exon 5 from *Impact* were amplified using PCR (primers detailed in Table 2.3) and analysed via Sanger sequencing. Unfortunately, no SNPs were observed in either exon 2 or exon 5 (Figure 3.13). Both exons contained the 129/sv variant, indicating that the *Impact* locus in the TET TKO line is only of 129/sv descent. Interestingly, this contrasts with information supplied by the creators of this line which states the TET TKO is a C57BL/6J x 129Sv cross (see GEO database for more information https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM1978027).

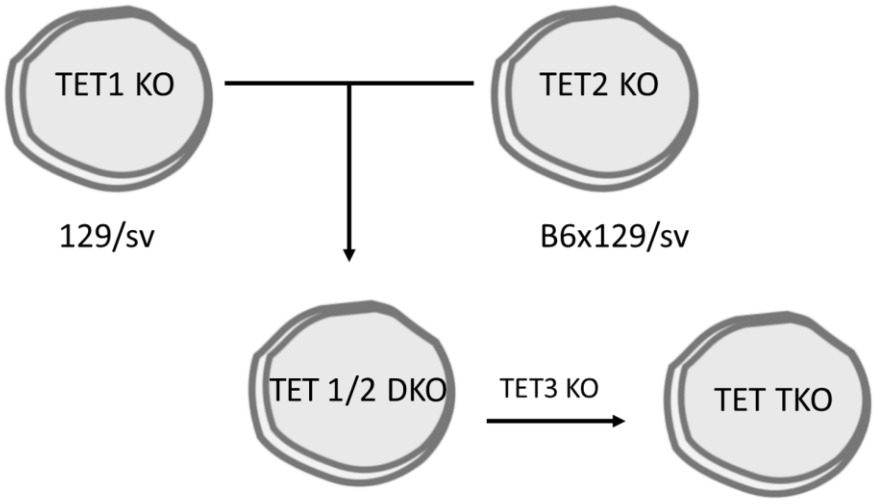


Figure 3.12 Derivation of the TET triple knockout cell line

Schematic diagram depicting the genetic background of the TET TKO cell line. Specifically, a TET1 knockout was performed in a 129/sv ES cell line, while a TET2 knockout was performed in a B6x129/sv ES cell line. The two knockout lines were crossed to produce a TET 1/2 double knockout, which then had TET3 knocked out to produce the TET triple knockout line (Hu et al., 2014).

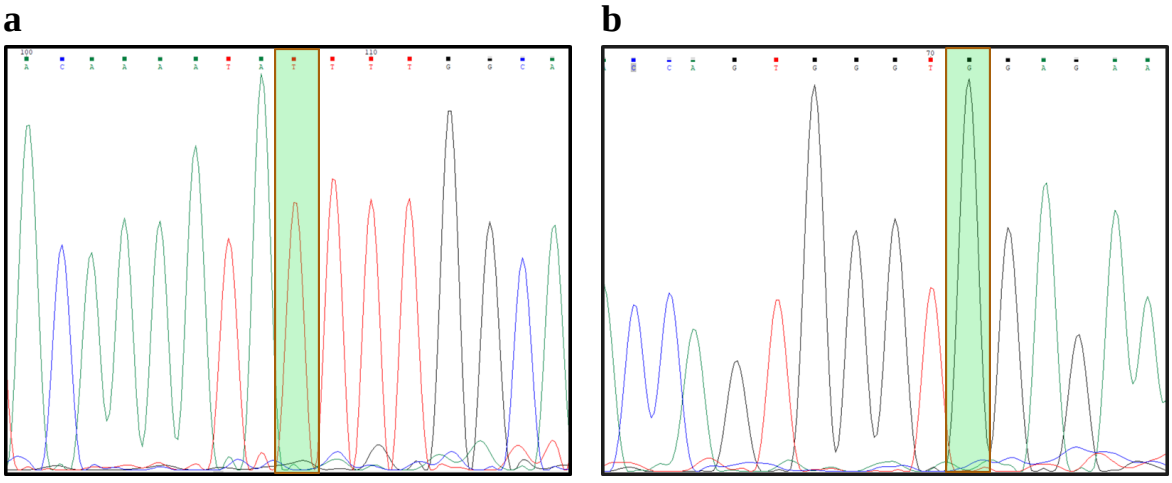


Figure 3.13 SNP analysis of the *Impact* locus in TET TKO cells

Chromatogram depicting the location of known SNPs between B6 and 129/sv mouse strains at **(a)** exon 2 and **(b)** exon 5 of the *Impact* locus. In both cases, the SNPs in question are clearly mono-allelic and 129/sv in origin.

3.3.3 TET1 overexpression and FACS analysis.

TET1 overexpression was performed by the addition of 1 μM doxycycline to culture media 72 hours before harvesting. To isolate and collect cells with either very low or very high TET1 expression fluorescence activated cell sorting was used. The following gating parameters were used:

- Samples were gated by FSC-Area v FSC-Height (Figure 3.14a) and SSC-Area and SSC-Height (Figure 3.14b) to remove double cells.
- Gating for cells with low DAPI expression ensured only live cells were selected (Figure 3.14c).
- Gating for cells with a 1:1 ratio of FSC-Area to SSC-Area selected cells based on size and morphology, excluding what is likely cellular debris (Figure 3.14d).
- Gating for cells with very low mCherry fluorescence (and therefore low TET1 expression) and cells with very high mCherry fluorescence (and therefore high TET1 expression), excluding those with moderate fluorescence (Figure 3.14e).

20,000 cells were collected for each antioxidant treatment range, 10,000 of each high and low TET1 expression.

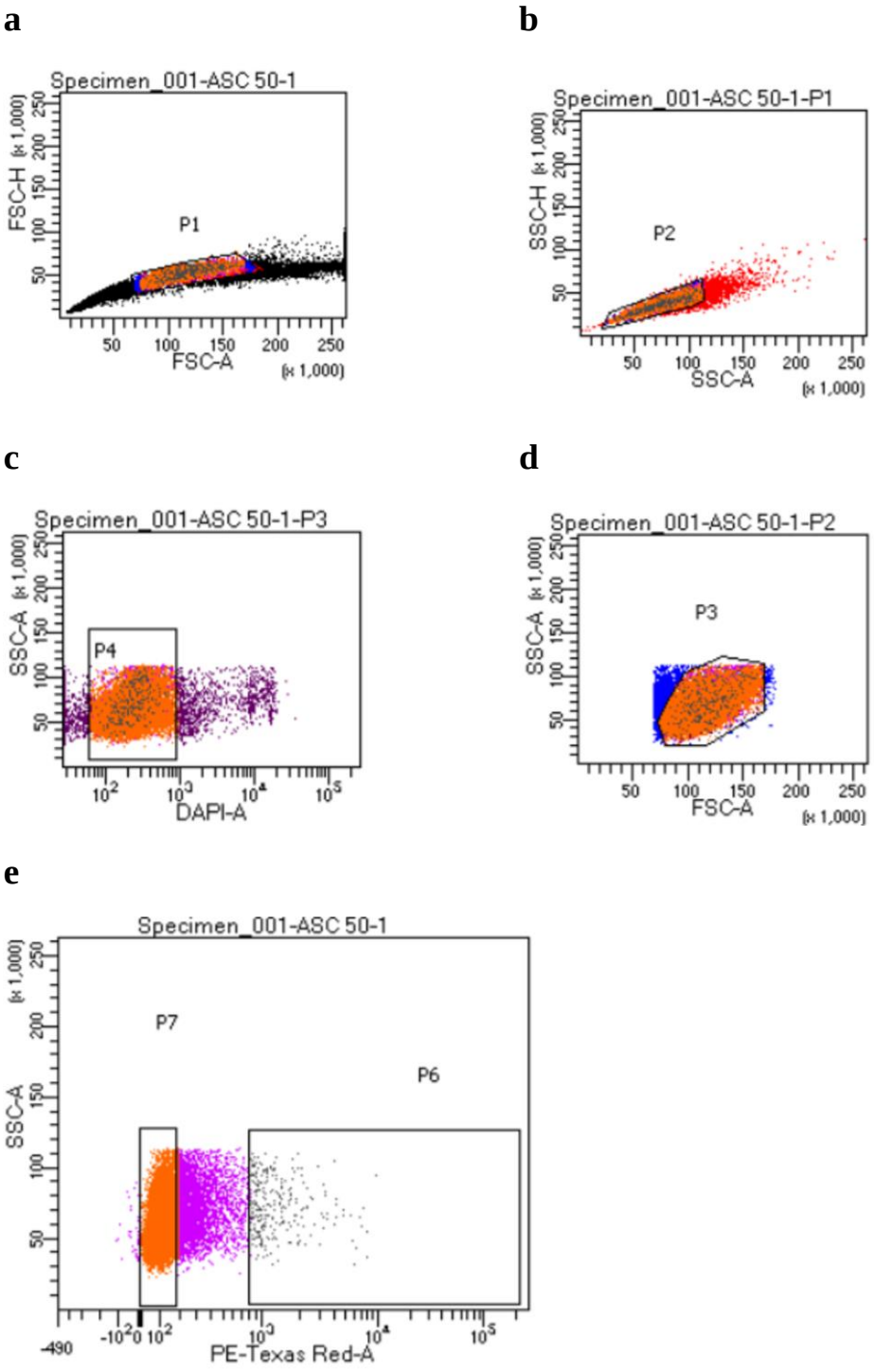


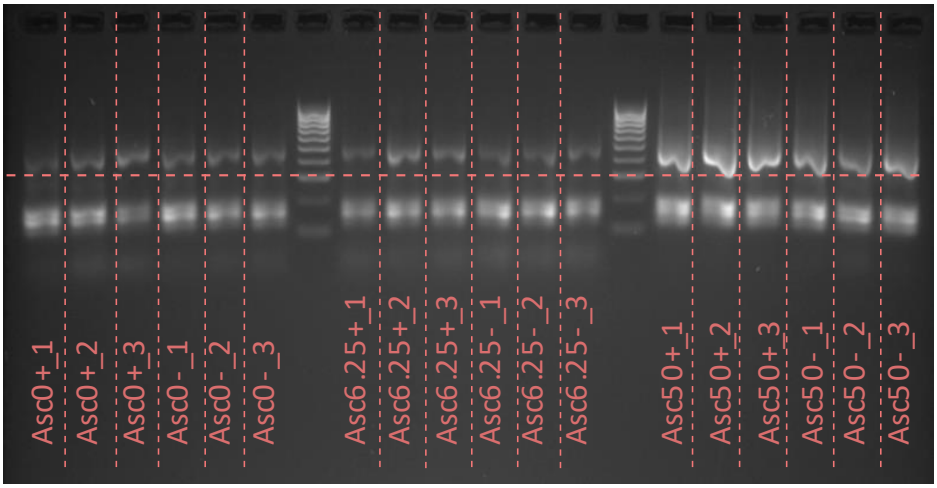
Figure 3.14 FACS of TET TKO cells after induction of TET1 expression.

Example of FACS gating after induction of TET1 expression in TET TKO cells. Cells were gated to remove **(a- b)** doublets **(c)** dead cells **(d)** cellular debris, and to select for **(e)** cells with very high or very low TET1 expression.

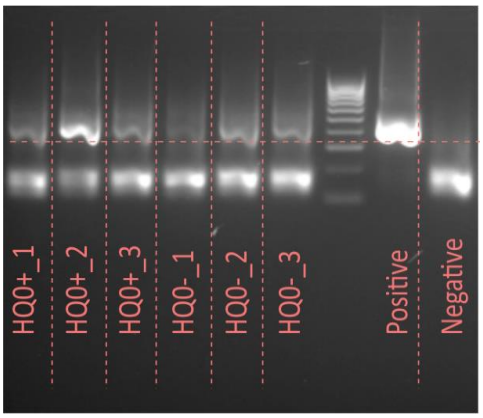
3.3.4 TET TKO - Bisulfite amplicon sequencing of KCNQ1ot1

Considering the *Impact* locus did not contain any allele specific SNPSs, and the KCNQ1ot1 PCR had to be optimised, KCNQ1ot1 was the locus of interest examined in this experiment. Using the PCR protocol described in 3.1.2, bisulfite amplicons were produced and sequenced via multiplex using the Illumina MiSeq system (Figure 3.15).

a



b



c

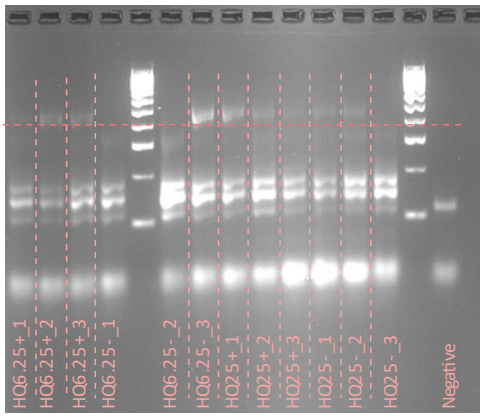


Figure 3.15 TET TKO ascorbate and hydroquinone amplicon library creation

Agarose gel electrophoresis analysis of amplicon libraries after cell treatment with either ascorbate or hydroquinone, n=3. 'Double bands' at 446 base pairs indicate KCNQ1ot1 amplicons with and without indexes attached. The three bands between 100 and 200 base pairs indicate of primer dimer.

The number of reads varied drastically between the samples (Figure 3.16). While some samples had more than 2000 reads, many did not surpass 500 reads. Considering this, it may be difficult to reach accurate conclusions from samples with extremely low read counts (i.e below 100).

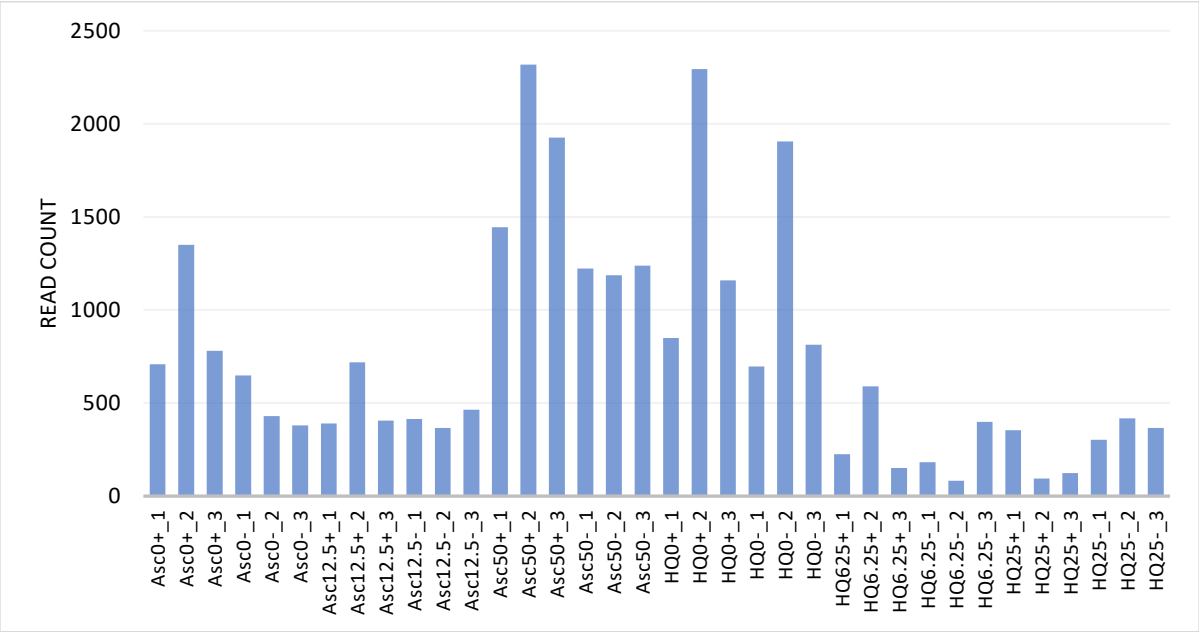


Figure 3.16 Read count from TET TKO mTET1-CD cells

Bar chart displaying the read count for all samples treated with ascorbate (Asc) and hydroquinone (HQ). + and – signs represent high and low TET expression respectively. The final concentration of supplemented Asc or HQ (ng/mL or µM respectively) is indicated for each sample.

3.3.4.1 Bisulfite conversion efficiency

Bisulfite conversion efficiency was determined using the same methodology as in section 3.2.2.1. All samples, with the exception of Asc50 High TET, have very high conversion efficiencies in line with those discussed previously. Asc50 High TET replicate 3 had a conversion efficiency of 95%, while the replicates 1 and 2 had

conversion efficiencies over 99.5%. Due to this, Asc50 High-TET_3 was removed from further analysis (Figure 3.17).

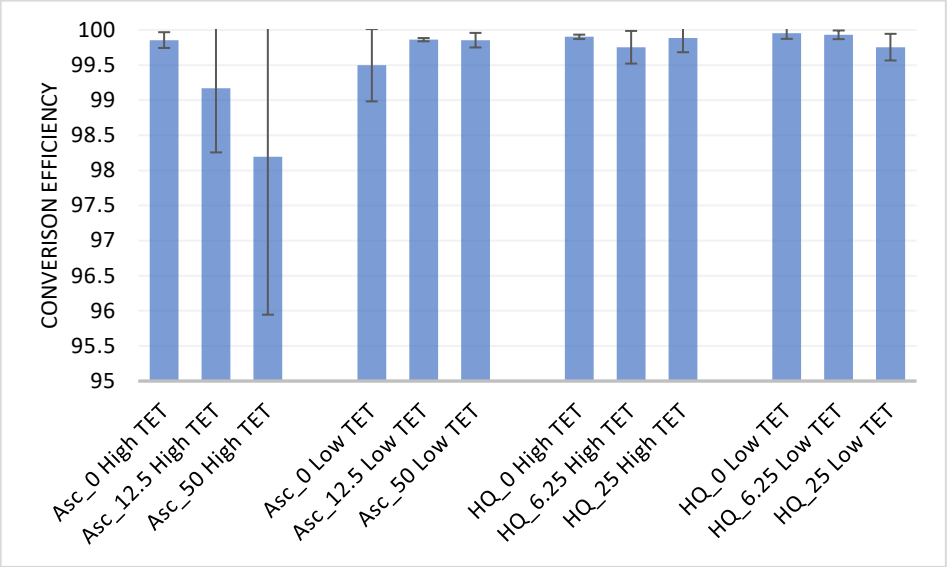


Figure 3.17 TET TKO mTET1-CD bisulfite conversion efficiency

The mean bisulfite conversion efficiency for ascorbate (Asc) and hydroquinone (HQ) treated samples, as determined by the ratio of unconverted to converted cytosines in the non-CpG context (n=3, ±S.D.). The final concentration of supplemented Asc or HQ (ng/mL or µM respectively) is indicated for each sample.

3.3.4.2 Methylation patterns

Despite high bisulfite conversion efficiency, all TET TKO samples presented with low read counts and repeated methylation patterns indicative of high levels of clonal expansion. Examples of these are given in supplementary Figure 5. This was enough to skew the data and prevent accurate conclusions from being drawn. Therefore, all TET TKO samples were omitted from the final results.

4 CHAPTER FOUR: DISCUSSION

Naïve ES cells and iPSCs have the potential to revolutionise the fields of regenerative medicine and mammalian transgenics. Artificial enhancement of TET enzyme activity via antioxidant supplementation can dramatically improve the reprogramming of differentiated cells to a naïve state of pluripotency (Coulter et al., 2013, Blaschke et al., 2013); however, the mechanism by which this occurs and its incidental effects are still being explored.

In this study, a potential negative consequence of TET-enhanced reprogramming was examined; specifically, its effect on genomic imprinting stability. This involved supplementing naïve ES cells with two different antioxidants, either ascorbate or hydroquinone, to boost TET activity through Fe^{2+} recycling within the TET catalytic domain. The methylation patterns at the KCNQ1ot1 locus were subsequently examined. Here, I have shown that ascorbate induces significant removal of DNA methylation. However, due to technical difficulties, I was unable to determine if this effect was also present during hydroquinone supplementation.

4.1 High throughput amplicon bisulfite sequencing

4.1.1 Advantages of multiplex bisulfite amplicon sequencing

To adequately answer the research question of this study the epigenetic status of a large number of samples needed to be analysed. To achieve this, I employed a high throughout multiplex sequencing approach similar to that recently used for HLA typing (Lange et al., 2014). However, what made my approach unique was (1) amplification

from bisulfite treated template, and (2) a novel two-step PCR whereby my region of interest was first amplified for 25 cycles, and then unique indexes complementary to the initial were added prior to another 10 cycles (Figure 2.2). The latter approach drastically reduced library preparation time compared with regular library construction. Double ended indexing was also a feature of the approach (i.e. having unique indexes at both the 5' and 3' end of the PCR amplicon), allowing potentially hundreds of different samples to be sequenced concurrently using the Illumina MiSeq system. Used together, these technical advances are considerably more powerful and faster than the plasmid vector cloning and Sanger sequencing techniques commonplace only a few years ago.

4.1.2 Limitations of multiplex bisulfite amplicon sequencing

Despite the positive features of multiplex bisulfite amplicon sequencing, there are several limitations that need to be addressed. The most prevalent of these during this study was clonal amplification in the TET TKO samples (Figure 3.9 and supplementary Figure 5). This occurs when a single sequence forms a disproportionate number of clusters compared to the other sequences within the same sample. When performing quantitative analyses, as was this case in the present study, clonal amplification makes any sequence affected appear more prevalent and thus skews the data. As a result, clonal amplification often renders data useless.

The biggest cause of clonal amplification is likely low template concentrations (Figure 3.6 and Figure 3.16). Therefore, improving initial template concentrations may be the best way to overcome this issue. While achieving higher cell numbers can be difficult when sorting cells with flow cytometry or FACS (for example, only 10,000 cells were gathered for each treatment during FACS), another solution may be a less

harsh bisulfite treatment. While this may reduce the overall bisulfite conversion efficiency, it may improve the DNA yield post conversion. This is because bisulfite conversion can lead to degradation of up to 90% of template DNA (Grunau et al., 2001).

However, it is possible that clonal amplification can still occur despite a higher amount of starting template. Therefore, it is also important not only to minimise the clonal amplification, but also be able to detect it when it is present. In this study, I identified clonal amplification by identification of amplification that favoured one allele over the other (Figure 3.8), and by analysis of the methylation patterns coming from the reads (Figure 3.9). While it was usually obvious when clonal amplification was present, the amount of diversity within each library was not formally quantified. In order to improve this, I would suggest using unique molecular identifiers within the gene-specific primers. This would involve the addition of a random 4-6 base pair sequence into the primers so that each amplicon would contain two indexes, one pertaining to the specific sample it came from, and another specific to that amplified molecule (Figure 4.1). This would allow any reads containing the same secondary index to be removed and give confidence that no read has been accounted for more than once.

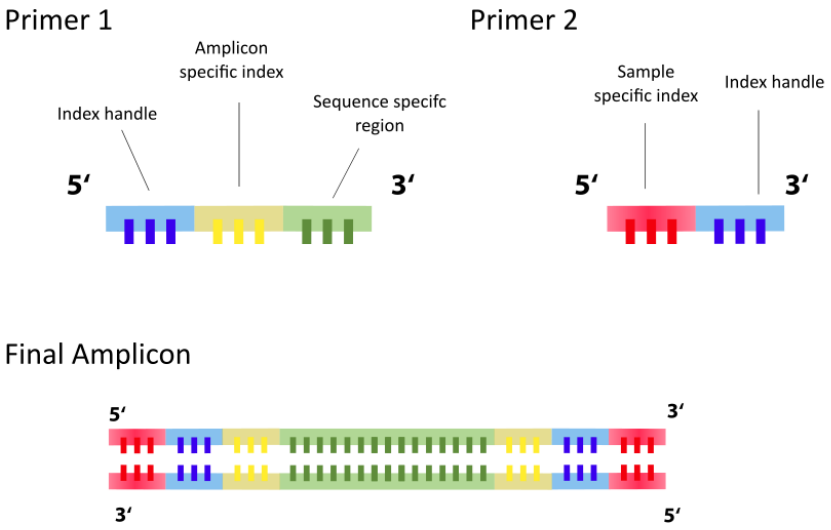


Figure 4.1 Addition of unique molecular indexes to PCR amplicons

Library generation would remain largely the same, however each amplicon would now have its own unique index to eliminate the effects of clonal expansion during sequencing.

4.2 Analysis of ascorbate treatment on KCNQ1ot1

4.2.1 The relationship between ascorbate, TET, and genomic imprinting stability

The purpose of this experiment was to determine if enhanced TET activity unintentionally results in the demethylation of the KCNQ1ot1 locus, subsequently resulting in loss of imprinting. I treated naïve B6xSD7 ES cells with various ascorbate concentrations ranging from 0 ng/µL to 400 ng/µL (Table 3.1), predicting that as the ascorbate concentration was increased there would be a corresponding, dose dependant loss of DNA methylation.

Interestingly, only samples that had been treated with 25 ng/µL had a statistically significant loss of methylation following correction for multiple testing; however, it appeared that samples between the ranges of 12.5 ng/µL to 50 ng/µL also lost

methylation that was nearing significance (i.e. $p<0.05$). In line with this, ICRs are resistant to the initial wave of demethylation following fertilization *in vivo* (Morgan et al., 2005), as well as the hypomethylation that occurs in naïve ES cells cultured in 2i. However, ICR demethylation does occur in 2i cultured cells, but much later than most other global demethylation events (Choi et al., 2017).

Surprisingly, loss of methylation did not appear to occur at 100 ng/μL of ascorbate; the highest concentration tested that did not show signs of clonal amplification. These results conform to the observations of Blaschke et al, (2013) who report no significant loss of methylation at ICRs after 100 ng/μL of ascorbate treatment. Interestingly, the aforementioned paper relied upon a single treatment concentration and concluded that enhanced TET activity through ascorbate supplementation had no significant effect on ICR stability. The results of the present study challenge this conclusion. Instead, I suggest that ascorbate induces a dose dependant effect on TET enzymatic activity where activity is highest in a 'sweet spot' of 12.5 ng/μL to 50 ng/μL. Therefore, I propose that TET-induced demethylation is modulated by ascorbate such that enzymatic activity increases to a peak level of activity at approximately 25 ng/μL of ascorbate, before subsequently lowering as ascorbate concentration increases (Figure 4.2). This coincides with the observation that knockout serum replacement (KSR), a common media used in ES cell and iPSC culture which contains ascorbate, induces increased aberrant methylation patterns at imprinted loci when compared to standard serum-based culture media (Lee et al., 2018).

While the PBAT technique was employed to assess global CpG methylation levels at each antioxidant concentration level, due to time constraints optimization was not completed in time and no results were gathered. However, Blaschke et al, (2013) have reported that ascorbate still improves TET-induced demethylation at 100 ng/μL on a

genome wide scale. Therefore, it is possible that any decrease in activity is minimal but still enough to be reduce activity levels below the threshold required for loss of methylation at ICRs.

This apparent decrease in TET activity during higher concentrations of ascorbate seems somewhat paradoxical, however I propose two hypotheses that I would like to test during future research in this subject area:

(1) Antioxidants such as ascorbate improve TET activity by enhancing Fe^{2+} recycling, but at higher concentrations inhibit the oxidation reaction of 5mC to 5hmC such that it is below the threshold for demethylation at ICRs.
(2) Ascorbate has a dual function as an antioxidant and a pro-oxidant in some contexts (Rietjens et al., 2002). At higher concentrations, this balance shifts towards a pro-oxidant state such that there is not enough Fe^{2+} to increase TET activity above the threshold for demethylation at ICRs.

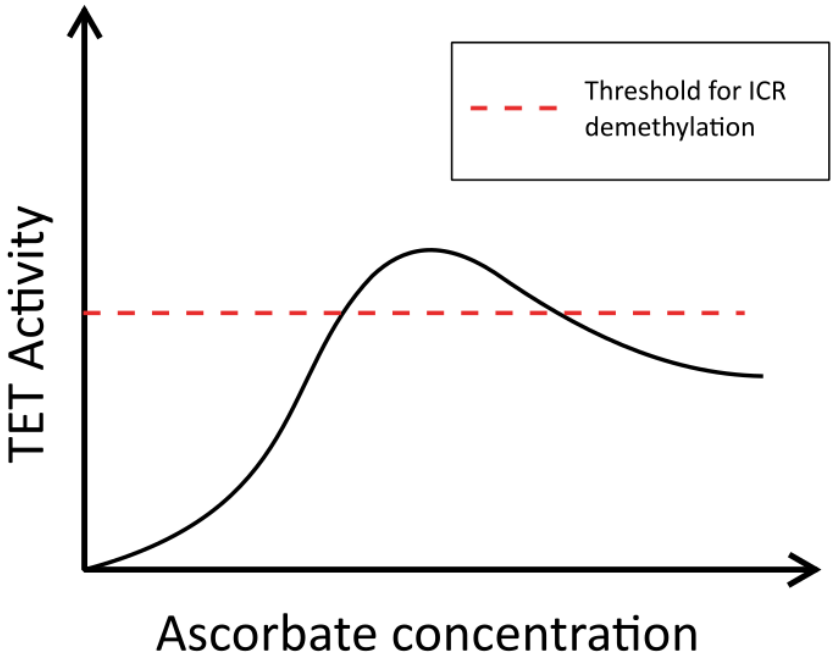


Figure 4.2 The relationship between ascorbate, TET, and imprinting stability

The proposed relationship between ascorbate supplementation, TET activity, and genomic imprinting stability in which there is a 'sweet spot' of enhanced activity that crosses a threshold and enables demethylation at imprint control regions.

4.2.2 Implications of the TET TKO experiment

Due to the TET TKO experiment suffering from clonal amplification, it is potentially difficult to say the effects observed in the B6xSD7 cell line under ascorbate treatment were due to TET activity alone. Ascorbate has many roles within the cell, and are likely to enhance the activity of other Fe^{2+} dependant dioxygenase enzymes. One class of these enzymes is the Jumanji (Jmj) histone demethylases, which remove methylation from histones by oxidation in a similar fashion to TET removal of methylated cytosine (Young et al., 2015, PMID: 25974700). Despite this, the overwhelming consensus from current literature is that global DNA methylation levels are directly impacted by ascorbate in a TET-dependent manner because removal of

TET abolishes ascorbate-induced global demethylation *(Blaschke et al., 2013, Minor et al., 2013).* Nevertheless, my research was not able to exclude the possibility that ascorbate has indirect effects on imprint stability which are independent of TET; an aspect of which should be further tested in subsequent experiments.

4.3 Other enhancers of TET

While I was unable to determine if hydroquinone supplementation induced TET-mediated demethylation at ICRs, the results from the ascorbate experiment show it can occur. As a result, any enhancer of TET may have the ability to cause loss of imprinting in pluripotent stem cells or IVF embryos, with the possibility of an additive effect across all components occurring. Perhaps the best example of this is retinoic acid (RA) which can work synergistically with ascorbate to improve TET activity and induce reprogramming. RA directly upregulates the expression of TET2 and TET3 by binding to a retinoic acid receptor element upstream of their promotors, causing subsequent transcription (Hore et al., 2016).

It is likely that due to the compounding action of these two vitamins that the imprinting instability reported in this study would be greatly amplified. Therefore, media design for future iPSC and IVF embryo culture should not consider the amount of ascorbate present within the media, but consider what effect the media composition as a whole will have on TET activity to ensure it does not pass the threshold for ICR demethylation.

4.4 Implications for regenerative medicine and *in vitro* Fertilization

The results of the current research have uncovered the potential danger of enhancing TET activity in pluripotent stem cells. In the introductory chapter of this thesis I discussed the consequences of aberrant methylation patterns at ICRs and the subsequent loss of imprinting. Specifically, it appears to be a driver for many cancers, in particular myeloid leukaemia (Jelinic and Shaw, 2007). So, while iPSCs are expected to be at the forefront of personalised regenerative medicine, and faster derivation of these cells will result in faster treatment for patients, it is a priority to maintain the health of these cells above all else. Any affected iPSCs will be unsafe for regenerative medicine applications. Therefore, if ascorbate-induced enhancement of TET is to be used during the reprograming of iPSCs then media conditions must be designed with this in mind, especially when considering given the relationship between LOI and myeloid leukaemia, and that one of the most common forms of stem cell therapy employed at the moment are haemopoietic stem cell transplantations (Passweg et al., 2012, Jelinic and Shaw, 2007).

Children born to assisted reproductive technologies (ART), such as *in vitro* fertilization (IVF), have a much greater risk of suffering from imprinting disorders, specifically Beckwith–Wiedemann syndrome (BWS) (Amor and Halliday, 2008). BWS is often caused by hypomethylation of several imprinted loci, including H19/IGF-2 and KCNQ1ot1, resulting in biallelic expression. This results in an increased risk of childhood cancer, as well as an array of congenital defects (Amor and Halliday, 2008, Chiesa et al., 2012). It is nine times more common in IVF conceived children (Halliday

et al., 2004), implying that people conceived via IVF are more prone to aberrant imprinting when compared to naturally conceived individuals.

Given that culture media is heavily deterministic of phenotype in ES cells, it is possible that the media used to culture IVF embryos may also have an impact on their epigenetic stability. Indeed, recently it was shown that porcine embryos cultured in standard, commercially available IVF media experienced greater aberrations in DNA methylation than in embryos cultured in reproductive fluid, or *in vivo* (Canovas et al., 2017). Unfortunately, because the media used to culture IVF embryos is becoming increasing commercialized, proprietary reasons have prevented the exact formulations of these media from being made public. However, some IVF culture media do contain ascorbic acid (Chronopoulou and Harper, 2015), however, it is often not known how much ascorbate is being added to media. Given the results from this study, it is possible that ascorbate-induced TET activity is responsible for the aberrant methylation observed in IVF embryos. Therefore, greater exploration into the effects of culture media, especially ascorbate and other antioxidants, on IVF embryos is required to improve culture media conditions and the overall health of those conceived from assisted reproductive technologies.

4.5 Concluding remarks

This study was part of an overall research project aiming to understand the properties of pluripotent stem cells and how to best harness their developmental potency. The use of small molecules that enhance erasure of epigenetic memory through the TET enzymes has drastically improved the rate at which differentiated cells can be reprogrammed back to the naïve state of pluripotency. However, the results of the ascorbate treatment

experiment show that this technique has downsides which effect the safety and viability of pluripotent cells in regenerative medicine. Despite the remainder of the study being hampered by technical difficulties, my results here can add to the current collection of literature describing the effects of culture media components on the phenotype of early mammalian embryos and ES cells. In doing so, this may allow the rational design of culture media that can improve both the efficiency and safety of naïve ESCs and embryos used in regenerative medicine, mammalian transgenics and IVF.

REFERENCES

ABERLE, H., BAUER, A., STAPPERT, J., KISPERT, A. & KEMLER, R. 1997. Beta-catenin is a target for the ubiquitin-proteasome pathway. *The EMBO Journal,* 16**,** 3797-3804.

AMOR, D. J. & HALLIDAY, J. 2008. A review of known imprinting syndromes and their association with assisted reproduction technologies. *Hum Reprod,* 23**,** 2826-34.

BLASCHKE, K., EBATA, K. T., KARIMI, M. M., ZEPEDA-MARTÍNEZ, J. A., GOYAL, P., MAHAPATRA, S., TAM, A., LAIRD, D. J., HIRST, M., RAO, A., LORINCZ, M. C. & RAMALHO-SANTOS, M. 2013. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature,* 500**,** 222.

CANOVAS, S., IVANOVA, E., ROMAR, R., GARCIA-MARTINEZ, S., SORIANO-UBEDA, C., GARCIA-VAZQUEZ, F. A., SAADEH, H., ANDREWS, S., KELSEY, G. & COY, P. 2017. DNA methylation and gene expression changes derived from assisted reproductive technologies can be decreased by reproductive fluids. *Elife,* 6.

CAPECCHI, M. R. 1989. The new mouse genetics: altering the genome by gene targeting. *Trends Genet,* 5**,** 70-6.

CHAMBERS, I., SILVA, J., COLBY, D., NICHOLS, J., NIJMEIJER, B., ROBERTSON, M., VRANA, J., JONES, K., GROTEWOLD, L. & SMITH, A. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature,* 450**,** 1230-4.

CHIESA, N., DE CRESCENZO, A., MISHRA, K., PERONE, L., CARELLA, M., PALUMBO, O., MUSSA, A., SPARAGO, A., CERRATO, F., RUSSO, S., LAPI, E., CUBELLIS, M. V., KANDURI, C., CIRILLO SILENGO, M., RICCIO, A. & FERRERO, G. B. 2012. The KCNQ1OT1 imprinting control region and non-coding RNA: new properties derived from the study of Beckwith-Wiedemann syndrome and Silver-Russell syndrome cases. *Hum Mol Genet,* 21**,** 10-25.

CHOI, J., HUEBNER, A. J., CLEMENT, K., WALSH, R. M., SAVOL, A., LIN, K., GU, H., DI STEFANO, B., BRUMBAUGH, J., KIM, S. Y., SHARIF, J., ROSE, C. M., MOHAMMAD, A., ODAJIMA, J., CHARRON, J., SHIODA, T., GNIRKE, A., GYGI, S., KOSEKI, H., SADREYEV, R. I., XIAO, A., MEISSNER, A. & HOCHEDLINGER, K. 2017. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature,* 548**,** 219-223.

CHRONOPOULOU, E. & HARPER, J. C. 2015. IVF culture media: past, present and future. *Human Reproduction Update,* 21**,** 39-55.

COSTA, Y., DING, J., THEUNISSEN, T. W., FAIOLA, F., HORE, T. A., SHLIAHA, P. V., FIDALGO, M., SAUNDERS, A., LAWRENCE, M., DIETMANN, S., DAS, S., LEVASSEUR, D. N., LI, Z., XU, M., REIK, W., SILVA, J. C. R. & WANG, J. 2013. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature,* 495**,** 370.

COULTER, J. B., O'DRISCOLL, C. M. & BRESSLER, J. P. 2013. Hydroquinone increases 5-hydroxymethylcytosine formation through ten eleven translocation 1 (TET1) 5-methylcytosine dioxygenase. *J Biol Chem,* 288**,** 28792-800.

DAWLATY, M. M., BREILING, A., LE, T., BARRASA, M. I., RADDATZ, G., GAO, Q., POWELL, B. E., CHENG, A. W., FAULL, K. F., LYKO, F. & JAENISCH, R. 2014. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Developmental cell,* 29**,** 102-111.

DOETSCHMAN, T., GREGG, R. G., MAEDA, N., HOOPER, M. L., MELTON, D. W., THOMPSON, S. & SMITHIES, O. 1987. Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature,* 330**,** 576.

EVANS, M. J. & KAUFMAN, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature,* 292**,** 154.

FICZ, G., HORE, T. A., SANTOS, F., LEE, H. J., DEAN, W., ARAND, J., KRUEGER, F., OXLEY, D., PAUL, Y. L., WALTER, J., COOK, S. J., ANDREWS, S., BRANCO, M. R. & REIK, W. 2013. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell,* 13**,** 351-9.

GAFNI, O., WEINBERGER, L., MANSOUR, A. A., MANOR, Y. S., CHOMSKY, E., BEN-YOSEF, D., KALMA, Y., VIUKOV, S., MAZA, I., ZVIRAN, A., RAIS, Y., SHIPONY, Z., MUKAMEL, Z., KRUPALNIK, V., ZERBIB, M., GEULA, S., CASPI, I., SCHNEIR, D., SHWARTZ, T., GILAD, S., AMANN-ZALCENSTEIN, D., BENJAMIN, S., AMIT, I., TANAY, A., MASSARWA, R., NOVERSHTERN, N. & HANNA, J. H. 2013. Derivation of novel human ground state naive pluripotent stem cells. *Nature,* 504**,** 282-6.

GARDNER, R. L. & BROOK, F. A. 1997. Reflections on the biology of embryonic stem (ES) cells. *The International journal of developmental biology,* 41**,** 235-243.

GOLDBERG, A. D., ALLIS, C. D. & BERNSTEIN, E. 2007. Epigenetics: A Landscape Takes Shape. *Cell,* 128**,** 635-638.

GRUNAU, C., CLARK, S. J. & ROSENTHAL, A. 2001. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Research,* 29**,** e65-e65.

GU, T. P., GUO, F., YANG, H., WU, H. P., XU, G. F., LIU, W., XIE, Z. G., SHI, L., HE, X., JIN, S. G., IQBAL, K., SHI, Y. G., DENG, Z., SZABO, P. E., PFEIFER, G. P., LI, J. & XU, G. L. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature,* 477**,** 606-10.

GUO, G., YANG, J., NICHOLS, J., HALL, J. S., EYRES, I., MANSFIELD, W. & SMITH, A. 2009. Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development,* 136**,** 1063-9.

HABIBI, E., BRINKMAN, A. B., ARAND, J., KROEZE, L. I., KERSTENS, H. H., MATARESE, F., LEPIKHOV, K., GUT, M., BRUN-HEATH, I., HUBNER, N. C., BENEDETTI, R., ALTUCCI, L., JANSEN, J. H., WALTER, J., GUT, I. G., MARKS, H. & STUNNENBERG, H. G. 2013. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell,* 13**,** 360-9.

HALLIDAY, J., OKE, K., BREHENY, S., ALGAR, E. & J. AMOR, D. 2004. Beckwith-Wiedemann Syndrome and IVF: A Case-Control Study. *American Journal of Human Genetics,* 75**,** 526-528.

HEMBERGER, M., DEAN, W. & REIK, W. 2009. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nature Reviews Molecular Cell Biology,* 10**,** 526.

HON, G. C., RAJAGOPAL, N., SHEN, Y., MCCLEARY, D. F., YUE, F., DANG, M. D. & REN, B. 2013. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat Genet,* 45**,** 1198-206.

HORE, T. A., VON MEYENN, F., RAVICHANDRAN, M., BACHMAN, M., FICZ, G., OXLEY, D., SANTOS, F., BALASUBRAMANIAN, S., JURKOWSKI, T. P. & REIK, W. 2016. Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naive pluripotency by complementary mechanisms. *Proc Natl Acad Sci U S A,* 113**,** 12202-12207.

HU, X., ZHANG, L., MAO, S. Q., LI, Z., CHEN, J., ZHANG, R. R., WU, H. P., GAO, J., GUO, F., LIU, W., XU, G. F., DAI, H. Q., SHI, Y. G., LI, X., HU, B., TANG, F., PEI, D. & XU, G. L. 2014. Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell,* 14**,** 512-22.

HUTTER, G., NOWAK, D., MOSSNER, M., GANEPOLA, S., MUSSIG, A., ALLERS, K., SCHNEIDER, T., HOFMANN, J., KUCHERER, C., BLAU, O., BLAU, I. W., HOFMANN, W. K. & THIEL, E. 2009. Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation. *N Engl J Med,* 360**,** 692-8.

IANNACCONE, P. M. & JACOB, H. J. 2009. Rats! *Disease Models & Mechanisms,* 2**,** 206-210.

ITO, S., D'ALESSIO, A. C., TARANOVA, O. V., HONG, K., SOWERS, L. C. & ZHANG, Y. 2010. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature,* 466**,** 1129-33.

JELINIC, P. & SHAW, P. 2007. Loss of imprinting and cancer. *J Pathol,* 211**,** 261-8.

JOHNSON, W. H., LOSKUTOFF, N. M., PLANTE, Y. & BETTERIDGE, K. J. 1995. Production of four identical calves by the separation of blastomeres from an in vitro derived four-cell embryo. *The Veterinary record,* 137**,** 15-16.

KALKAN, T. & SMITH, A. 2014. Mapping the route from naive pluripotency to lineage specification. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 369**,** 20130540.

KIM, M., YAN, Y., LEE, K., SGAGIAS, M. & COWAN, K. H. 2004. Ectopic expression of von Hippel-Lindau tumor suppressor induces apoptosis in 786-O renal cell carcinoma cells and regresses tumor growth of 786-O cells in nude mouse. *Biochem Biophys Res Commun,* 320**,** 945-50.

KOLLER, B. H., MARRACK, P., KAPPLER, J. W. & SMITHIES, O. 1990. Normal development of mice deficient in beta 2M, MHC class I proteins, and CD8+ T cells. *Science,* 248**,** 1227-30.

KRUEGER, C., KING, M. R., KRUEGER, F., BRANCO, M. R., OSBORNE, C. S., NIAKAN, K. K., HIGGINS, M. J. & REIK, W. 2012. Pairing of Homologous Regions in the mouse genome is associated with transcription but not imprinting status. *PLOS ONE,* 7**,** e38983.

KUEHN, M. R., BRADLEY, A., ROBERTSON, E. J. & EVANS, M. J. 1987. A potential animal model for Lesch-Nyhan syndrome through introduction of HPRT mutations into mice. *Nature,* 326**,** 295-8.

KUNATH, T., SABA-EL-LEIL, M. K., ALMOUSAILLEAKH, M., WRAY, J., MELOCHE, S. & SMITH, A. 2007. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development,* 134**,** 2895-902.

LANGE, V., BOHME, I., HOFMANN, J., LANG, K., SAUTER, J., SCHONE, B., PAUL, P., ALBRECHT, V., ANDREAS, J. M., BAIER, D. M., NETHING, J., EHNINGER, U., SCHWARZELT, C., PINGEL, J., EHNINGER, G. & SCHMIDT, A. H. 2014. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC Genomics,* 15**,** 63.

LEE, E. C., LIANG, Q., ALI, H., BAYLISS, L., BEASLEY, A., BLOOMFIELD-GERDES, T., BONOLI, L., BROWN, R., CAMPBELL, J., CARPENTER, A., CHALK, S., DAVIS, A., ENGLAND, N., FANE-DREMUCHEVA, A., FRANZ, B., GERMASCHEWSKI, V., HOLMES, H., HOLMES, S., KIRBY, I., KOSMAC, M., LEGENT, A., LUI, H., MANIN, A., O'LEARY, S., PATERSON, J., SCIARRILLO, R., SPEAK, A., SPENSBERGER, D., TUFFERY, L., WADDELL, N., WANG, W., WELLS, S., WONG, V., WOOD, A., OWEN, M. J., FRIEDRICH, G. A. & BRADLEY, A. 2014a. Complete humanization of the mouse immunoglobulin loci enables efficient therapeutic antibody discovery. *Nature Biotechnology,* 32**,** 356.

LEE, H. J., HORE, T. A. & REIK, W. 2014b. Reprogramming the methylome: erasing memory and creating diversity. *Cell Stem Cell,* 14**,** 710-9.

LEE, J., MATSUZAWA, A., SHIURA, H., SUTANI, A. & ISHINO, F. 2018. Preferable in vitro condition for maintaining faithful DNA methylation imprinting in mouse embryonic stem cells. *Genes Cells*.

LEE, J. H., PARK, S. J. & NAKAI, K. 2017. Differential landscape of non-CpG methylation in embryonic stem cells and neurons caused by DNMT3s. *Sci Rep,* 7**,** 11295.

LI, P., TONG, C., MEHRIAN-SHAI, R., JIA, L., WU, N., YAN, Y., MAXSON, R. E., SCHULZE, E. N., SONG, H., HSIEH, C. L., PERA, M. F. & YING, Q. L. 2008. Germline competent embryonic stem cells derived from rat blastocysts. *Cell,* 135**,** 1299-310.

LIU, Y., XU, H. W., WANG, L., LI, S. Y., ZHAO, C. J., HAO, J., LI, Q. Y., ZHAO, T. T., WU, W., WANG, Y., ZHOU, Q., QIAN, C., WANG, L. & YIN, Z. Q. 2018. Human embryonic stem cell-derived retinal pigment epithelium transplants as a potential treatment for wet age-related macular degeneration. *Cell Discovery,* 4**,** 50.

LU, F. & ZHANG, Y. 2015. Cell totipotency: molecular features, induction, and maintenance. *Natl Sci Rev,* 2**,** 217-225.

MARTELLO, G. & SMITH, A. 2014. The nature of embryonic stem cells. *Annu Rev Cell Dev Biol,* 30**,** 647-75.

MARTELLO, G., SUGIMOTO, T., DIAMANTI, E., JOSHI, A., HANNAH, R., OHTSUKA, S., GÖTTGENS, B., NIWA, H. & SMITH, A. 2012. Esrrb Is a Pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. *Cell Stem Cell,* 11**,** 491-504.

MINOR, E. A., COURT, B. L., YOUNG, J. I. & WANG, G. 2013. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem,* 288**,** 13669-74.

MIURA, F., ENOMOTO, Y., DAIRIKI, R. & ITO, T. 2012. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res,* 40**,** e136.

MORGAN, H. D., SANTOS, F., GREEN, K., DEAN, W. & REIK, W. 2005. Epigenetic reprogramming in mammals. *Hum Mol Genet,* 14 Spec No 1**,** R47-58.

MORISON, I. M., RAMSAY, J. P. & SPENCER, H. G. 2005. A census of mammalian imprinting. *Trends Genet,* 21**,** 457-65.

MORRIS, C. R., SINGER, S. T. & WALTERS, M. C. 2006. Clinical hemoglobinopathies: iron, lungs and new blood. *Curr Opin Hematol,* 13**,** 407-18.

NICHOLS, J. & SMITH, A. 2009. Naive and Primed Pluripotent States. *Cell Stem Cell,* 4**,** 487-492.

NICHOLS, J., ZEVNIK, B., ANASTASSIADIS, K., NIWA, H., KLEWE-NEBENIUS, D., CHAMBERS, I., SCHÖLER, H. & SMITH, A. 1998. Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct4. *Cell,* 95**,** 379-391.

OBERACKER, P., STEPPER, P., BOND, D. M., HOHN, S., FOCKEN, J., MEYER, V., SCHELLE, L., SUGRUE, V. J., JEUNEN, G.-J., MOSER, T., HORE, S. R., VON MEYENN, F., HIPP, K., HORE, T. A. & JURKOWSKI, T. P. 2018. Bio-On-Magnetic-Beads (BOMB): Open platform for high-throughput nucleic acid extraction and manipulation. *bioRxiv*.

PALMITER, R. D. & BRINSTER, R. L. 1986. Germ-Line Transformation of Mice. *Annual review of genetics,* 20**,** 465-499.

PASSWEG, J. R., BALDOMERO, H., GRATWOHL, A., BREGNI, M., CESARO, S., DREGER, P., DE WITTE, T., FARGE-BANCEL, D., GASPAR, B., MARSH, J., MOHTY, M., PETERS, C., TICHELLI, A., VELARDI, A., DE ELVIRA, C. R., FALKENBURG, F., SUREDA, A. & MADRIGAL, A. 2012. The EBMT activity survey: 1990-2010. *Bone Marrow Transplant,* 47**,** 906-23.

PLANELLO, A. C., JI, J., SHARMA, V., SINGHANIA, R., MBABAALI, F., MÜLLER, F., ALFARO, J. A., BOCK, C., DE CARVALHO, D. D. & BATADA, N. N. 2014. Aberrant DNA methylation reprogramming during induced pluripotent stem cell generation is dependent on the choice of reprogramming factors. *Cell Regeneration,* 3**,** 3:4.

RAMSAHOYE, B. H., BINISZKIEWICZ, D., LYKO, F., CLARK, V., BIRD, A. P. & JAENISCH, R. 2000. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A,* 97**,** 5237-42.

RIDEOUT, W. M., 3RD, WAKAYAMA, T., WUTZ, A., EGGAN, K., JACKSON-GRUSBY, L., DAUSMAN, J., YANAGIMACHI, R. & JAENISCH, R. 2000. Generation of mice from wild-type and targeted ES cells by nuclear cloning. *Nat Genet,* 24**,** 109-10.

RIETJENS, I. M. C. M., BOERSMA, M. G., HAAN, L. D., SPENKELINK, B., AWAD, H. M., CNUBBEN, N. H. P., VAN ZANDEN, J. J., WOUDE, H. V. D., ALINK, G. M. & KOEMAN, J. H. 2002. The pro-oxidant chemistry of the natural antioxidants vitamin C, vitamin E, carotenoids and flavonoids. *Environmental Toxicology and Pharmacology,* 11**,** 321-333.

ROYO, H. & CAVAILLE, J. 2008. Non-coding RNAs in imprinted gene clusters. *Biol Cell,* 100**,** 149-66.

SCHMIDT, C. S., BULTMANN, S., MEILINGER, D., ZACHER, B., TRESCH, A., MAIER, K. C., PETER, C., MARTIN, D. E., LEONHARDT, H. & SPADA, F. 2012. Global DNA hypomethylation prevents consolidation of differentiation programs and allows reversion to the embryonic stem cell state. *PLoS One,* 7**,** e52629.

SMALLWOOD, S. A., TOMIZAWA, S.-I., KRUEGER, F., RUF, N., CARLI, N., SEGONDS-PICHON, A., SATO, S., HATA, K., ANDREWS, S. R. & KELSEY, G. 2011. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature Genetics,* 43**,** 811.

SMITH, A. G., HEATH, J. K., DONALDSON, D. D., WONG, G. G., MOREAU, J., STAHL, M. & ROGERS, D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature,* 336**,** 688.

SOLTER, D. 2006. From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. *Nat Rev Genet,* 7**,** 319-27.

TAHILIANI, M., KOH, K. P., SHEN, Y., PASTOR, W. A., BANDUKWALA, H., BRUDNO, Y., AGARWAL, S., IYER, L. M., LIU, D. R., ARAVIND, L. & RAO, A. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* 324**,** 930-5.

TAKAHASHI, K. & YAMANAKA, S. 2006. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell,* 126**,** 663-676.

THEUNISSEN, T. W., POWELL, B. E., WANG, H., MITALIPOVA, M., FADDAH, D. A., REDDY, J., FAN, Z. P., MAETZEL, D., GANZ, K., SHI, L., LUNGJANGWA, T., IMSOONTHORNRUKSA, S., STELZER, Y., RANGARAJAN, S., D'ALESSIO, A., ZHANG, J., GAO, Q., DAWLATY, M. M., YOUNG, R. A., GRAY, N. S. & JAENISCH, R. 2014. Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell,* 15**,** 471-487.

VEIGA, A., CALDERON, G., BARRI, P. N. & COROLEU, B. 1987. Pregnancy after the replacement of a frozen-thawed embryo with less than 50% intact blastomeres. *Hum Reprod,* 2**,** 321-3.

VON MEYENN, F., IURLARO, M., HABIBI, E., LIU, N. Q., SALEHZADEH-YAZDI, A., SANTOS, F., PETRINI, E., MILAGRE, I., YU, M., XIE, Z., KROEZE, L. I., NESTEROVA, T. B., JANSEN, J. H., XIE, H., HE, C., REIK, W. & STUNNENBERG, H. G. 2016. Impairment of DNA Methylation Maintenance Is the Main Cause of Global Demethylation in Naive Embryonic Stem Cells. *Mol Cell,* 62**,** 848-861.

WILLIAMS, R. L., HILTON, D. J., PEASE, S., WILLSON, T. A., STEWART, C. L., GEARING, D. P., WAGNER, E. F., METCALF, D., NICOLA, N. A. & GOUGH, N. M. 1988. Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature,* 336**,** 684-7.

WRAY, J., KALKAN, T., GOMEZ-LOPEZ, S., ECKARDT, D., COOK, A., KEMLER, R. & SMITH, A. 2011. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nature Cell Biology,* 13**,** 838.

WRAY, J., KALKAN, T. & SMITH, A. G. 2010. The ground state of pluripotency. *Biochem Soc Trans,* 38**,** 1027-32.

YING, Q. L., WRAY, J., NICHOLS, J., BATLLE-MORERA, L., DOBLE, B., WOODGETT, J., COHEN, P. & SMITH, A. 2008. The ground state of embryonic stem cell self-renewal. *Nature,* 453**,** 519-23.

YOUNAN, P., KOWALSKI, J. & KIEM, H.-P. 2013. Genetic Modification of Hematopoietic stem cells as a therapy for HIV/AIDS. *Viruses,* 5**,** 2946-2962.

APPENDICES

Supplementary Figures

a

>Maternal (B6) KCNQ1ot1 amplicon sequence

GGCCTCAAGACCACCCCTGCTTCTGTAAGCCTGGGCCACAAAGAT----------GGGG

ACGTGGACGCAAAATACGAGAACTGAGCCACGGCCGTGAAACGAGGACCGGCCGTGAAACGA

GGACCGAGCCGTAACTGCAAAACGAATACGGAGCCACTGCGGCAAAACGAAGATGGAGCCCA

GCCGCGAAAGCGCGGCACGAATCACCTCTGCTTCTGGCCGTGAGTGCTTGCCGCGAGGAGGG

GGAGGCTATGATGAGCGCGGCCACGCGGACTTGCGACTTGTGCCGTGCTGACTCAGAGAAGA

AACCCGCGCTGAGAAAAAAACCATACCTAGGAGAACCATGCCGAGAAAAAGAAGCGCTGGGA

ACCAAGCTGAACAGAAA

b

>Paternal (SD7) KCNQ1ot1 amplicon sequence

GGCCTCAAGACCACCCCTGCTTCTGTAAGCCTGGGCCACAAAGATGGAGTTGTGGTTGGGGA

CGTGGACGCAAAATACGAGAACTGAGCCA----------CGGCCGTGAAACGAG

GACCGAGCCGTAACTGCAAAACGAATACGGAGCCACTGCGGCAAAACGAAGATGGAGCCCAG

CCGCGAAAGCGCGGCACGAATCACCTCTGCTTCTGGCCGTGAGTGCTTGCCGCGAGGAGGGG

GAGGCTATGATGAGCGCGGCCACGCGGACTTGCGACTTGTGCCGTGCTGACTCAGAGAAGAA

ACCCGCGCTGAGAAAAAAACCATACCTAGGAGAACCATGCCGAGAAAAAGAAGCGCTGGGAA

CCAAGCTGAACAGAAA

Supplementary Figure 1 Full KCNQ1ot1 amplicon sequence

The KCNQ1ot1 amplicon nucleotide sequence. **(a)** Maternal (C57BL/6) allele **(b)** Paternal (129/sv) allele

a

>*Impact* exon 2

GGCTGAAGAGGAAGTAGGGAACAGCCAGAGGCAGAGTGAAGAAATCGAAGCAATGGC

AGCCATTTATGGCGAGGAGTGGTGTGTCATTGATGA[**A/G**]AATGCCAAAATATTTT

GTATTAGAGTCACTGACTTCATGGATGACCCCAAATGGACACTTTGTTTACAGGTGA

TGTTGCCAAGTGAGTACCCGGGTACAGCACCGCCTTCTTATCA

b

>*Impact* exon 5

TACAGCACCGCCTTCTTATCAGCTGAACGCTCCCTGGCTGAAAGGGCAAGAACGCGC

AGACTTATCGAACAGCCTTGAGGAGATATATGTCCACAACATGGGTGAATCTATTCT

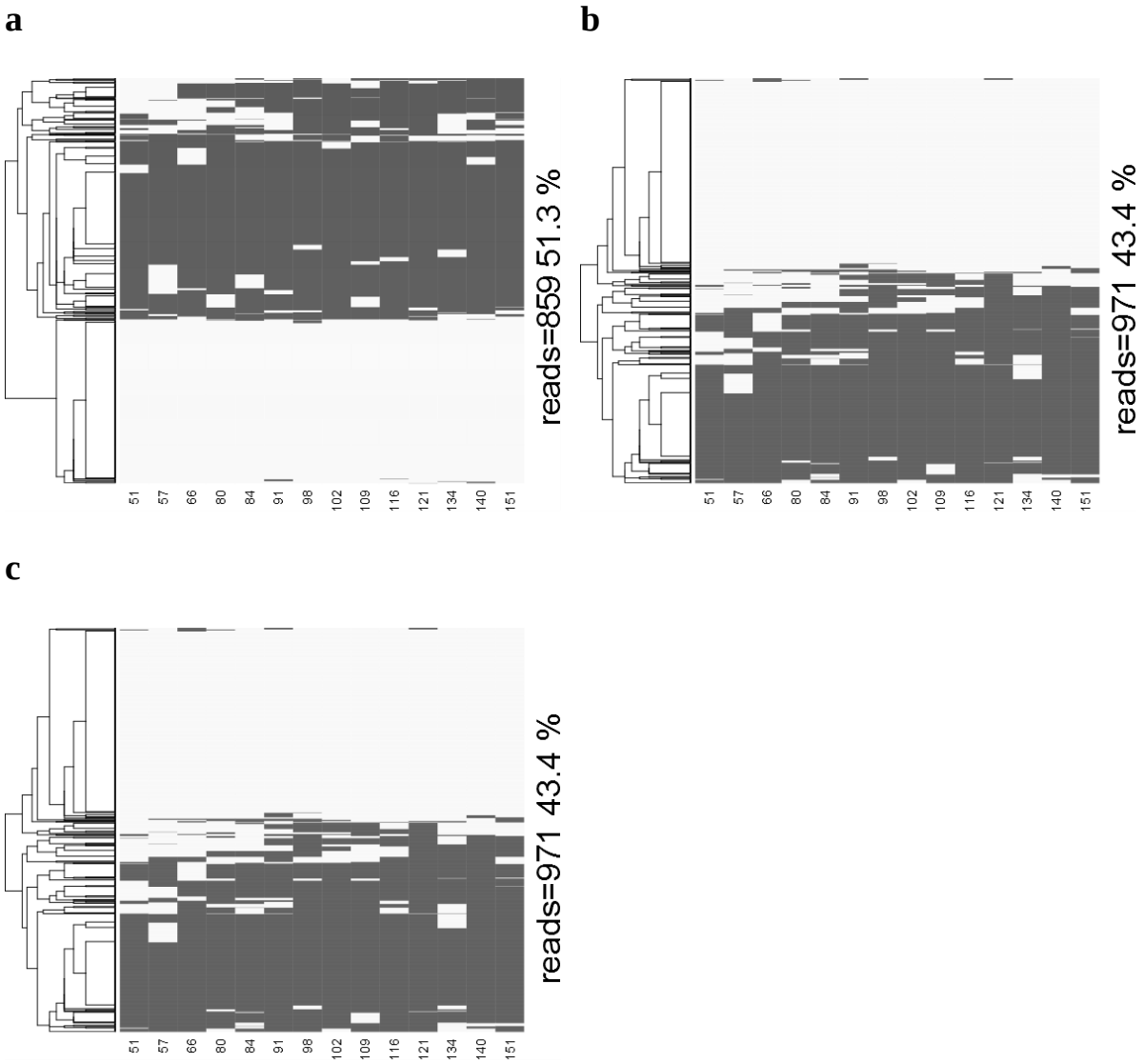
TTACCAGTGGGT[**G/C**]GAGAAAATAAGAGATGCTCTGATACAGAAATCTCAGATAA

CCGAGCCAGACCCAGATGTCAAGAAGAAAACTGAAGAGGTTGAGGTTGAGTCCGAAG

AAGAC

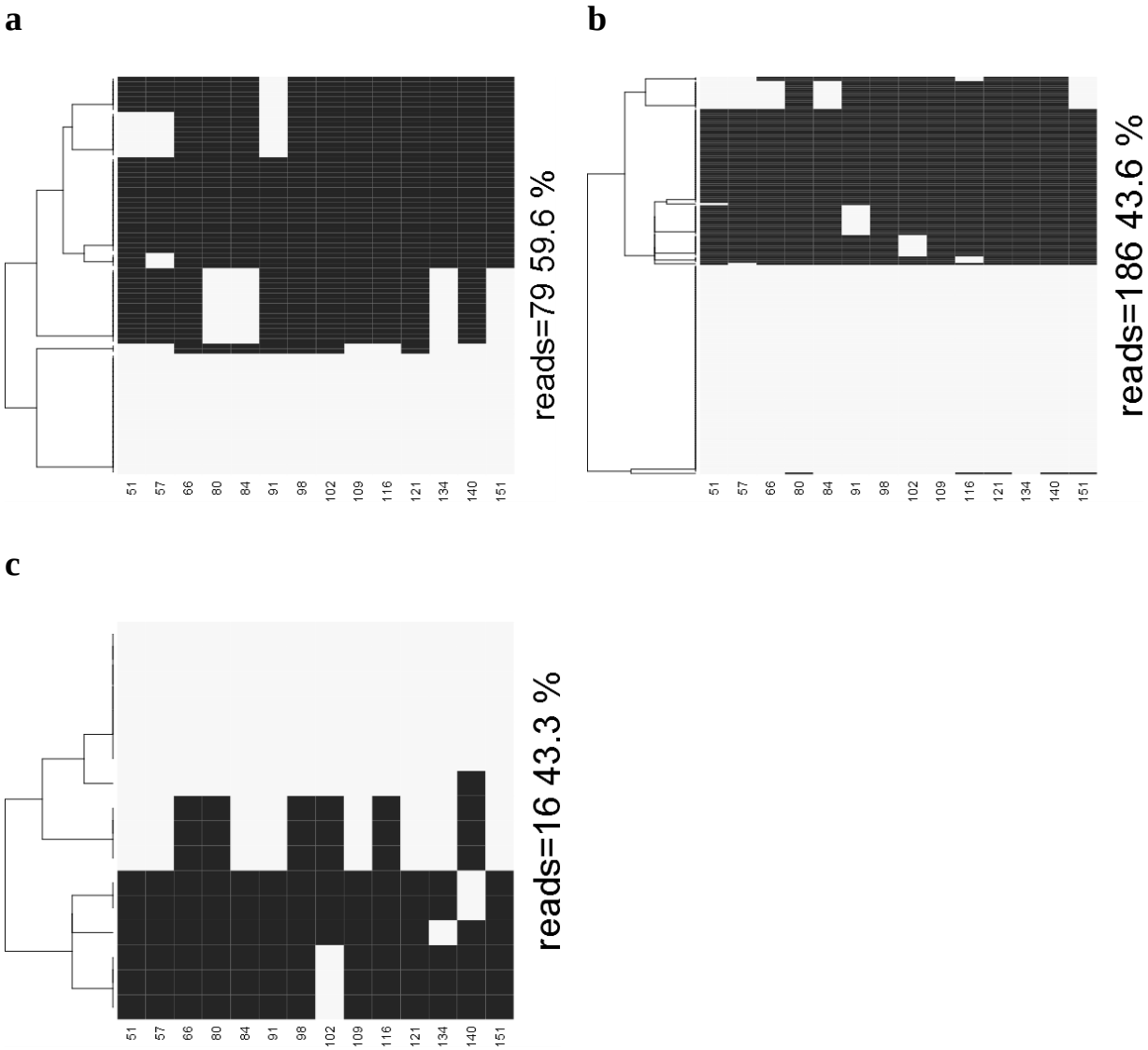
Supplementary Figure 2 Full Impact amplicon sequences

The *Impact* cDNA nucleotide sequence for **(a)** exon 2, and **(b)** exon 5. SNPs are highlighted within square brackets and formatted as [**129/sv / C57BL/6**].



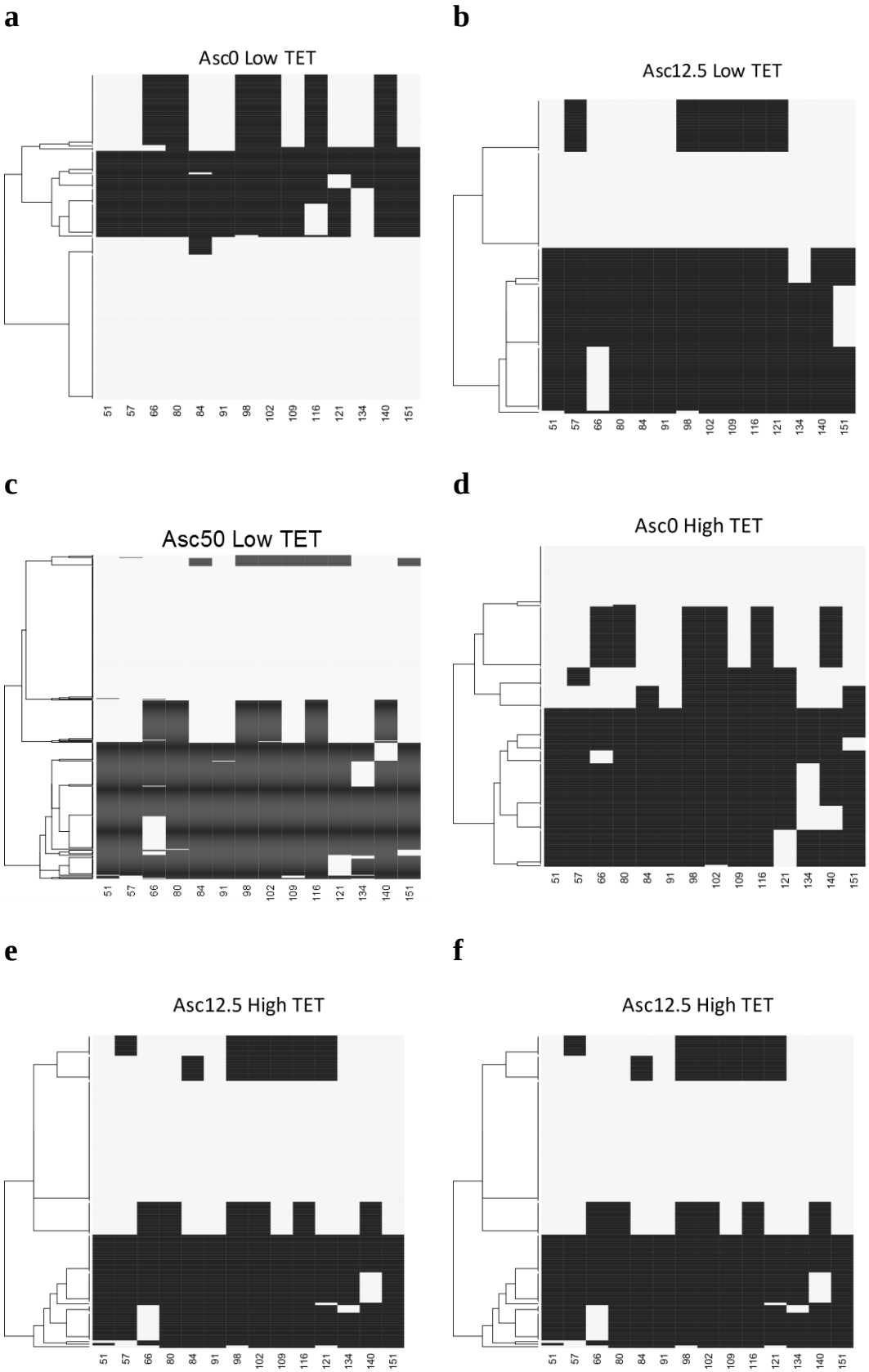
Supplementary Figure 3 KCNQ1ot1 methylation patterns from ascorbate treated B6xSD7 cells

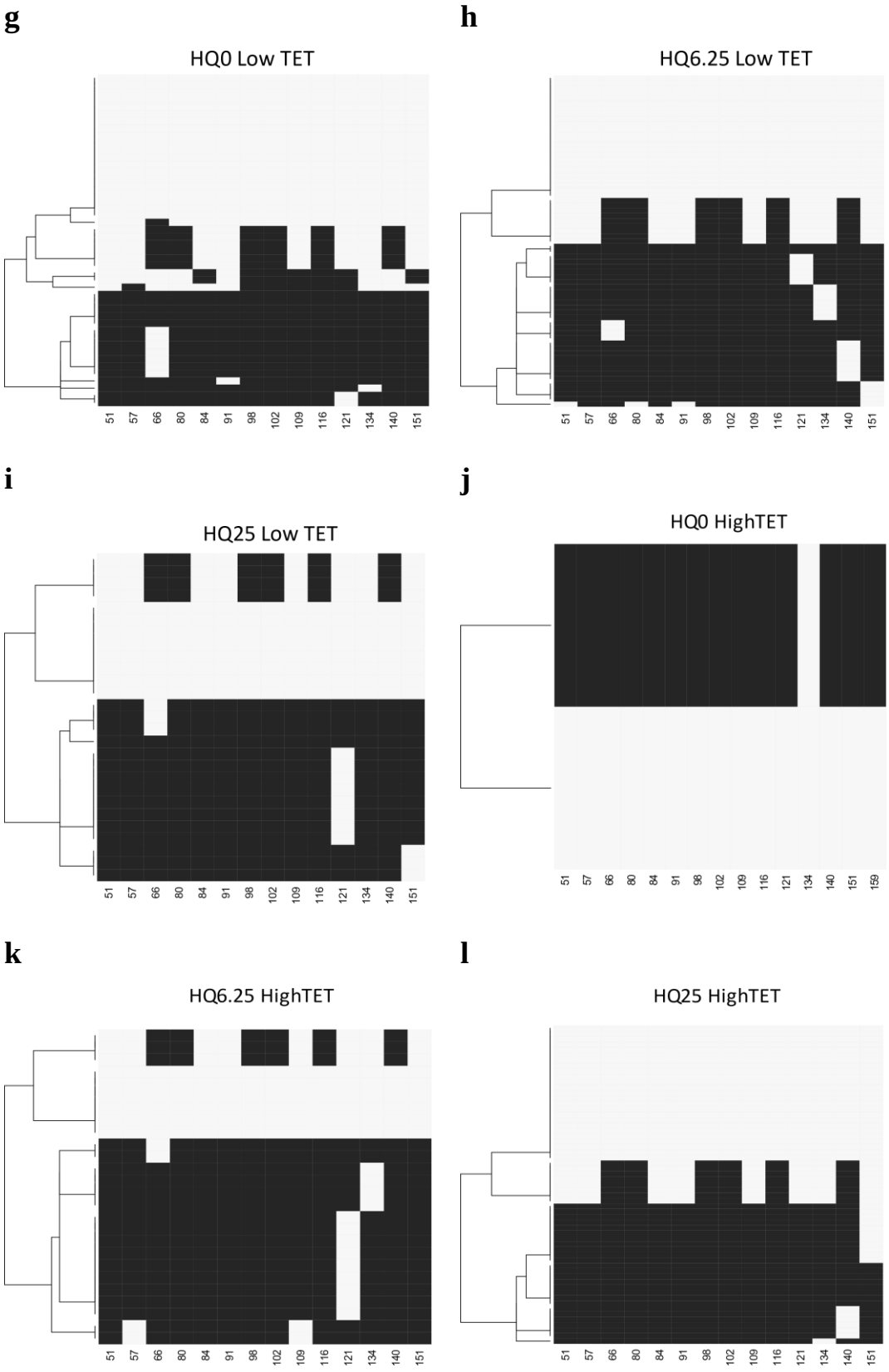
Examples of methylation patterns from B6xSD7 cells treated with **(a)** 0 ng/µL, **(b)** 25 ng/µL, or **(c)** 100 ng/µL of ascorbate



Supplementary Figure 4 KCNQ1ot1 methylation patterns from hydroquinone treated B6xSD7 cells

Examples of methylation patterns from B6xSD7 cells treated with **(a)** 0 µM, **(b)** 25 µM, or **(c)** 100 µM of hydroquinone

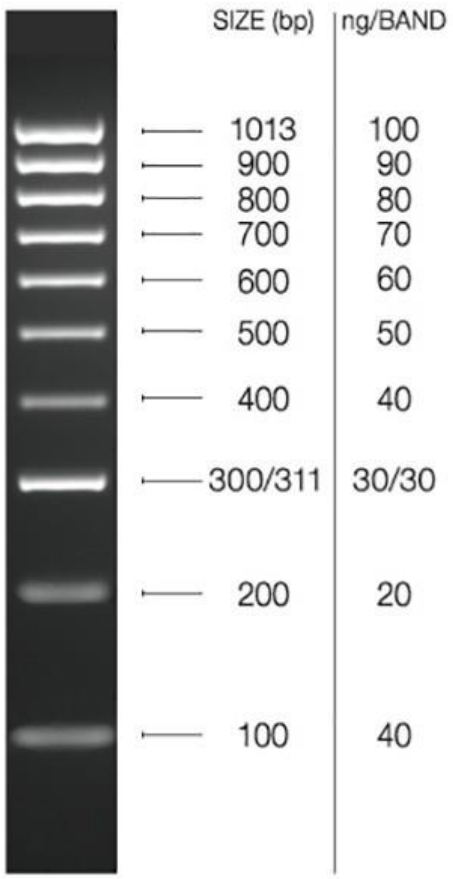

a
Asc0 Low TET
b
Asc12.5 Low TET
c
Asc50 Low TET
d
Asc0 High TET
e
Asc12.5 High TET
f
Asc12.5 High TET
51 57 66 80 84 91 98 102 109 116 121 134 140 151



Supplementary Figure 5 KCNQ1ot1 methylation patterns from TET TKO cells

Supplementary Figure 5 KCNQ1ot1 methylation patterns from TET TKO cells

Examples of methylation patterns from TET TKO cells treated with either ascorbate (Asc) or hydroquinone (HQ). The final concentration of supplemented Asc or HQ (ng/mL or μM respectively) is indicated for each sample. Figures specified as low TET were from cells with low mCherry fluorescence, while figures specified as high TET has high mCherry fluorescence during FACS.



Supplementary Figure 6 Bioline HyperLadder™ molecular weight maker sizes

Bioline HyperLadder™ 100bp

Supplementary Tables

Supplementary Table 1 FBS-based ESM recipe

FBS-based ESM Recipe	
Components	**Volume**
DMEM, high glucose, pyruvate (Gibco, 11995-065)	215 mL
FBS (Moregate Biotech FBS sterile filtered, MG-FBSF-500ML)	37.5mL
Penicillin-Streptomycin (100X, Life Tech, 15140-122)	2.5mL
MEM Non-Essential Amino Acids Solution (100X, Life Tech, 11140-050)	2.5mL
2-Mercaptoethanol (55 mM, Life Tech, 21985-023)	0.25mL
Total	257.75mL
LIF 10ug/mL (1000x, diluted in PBS + 0.1% BSA. Sourced from CSCR, Sally Lees)	1uL per mL of medium

Supplementary Table 2 N2B27 ESM recipe

N2B27 ESM Recipe	
Components	**Volume**
DMEM/F12 (Gibco 21331-020)	100mL
Neurobasal (Gibco 21103-049)	100mL
N2 (Stem Cells SF-NS-01-005)	1mL
B27 (Gibco 17504-044)	2mL
L-Glutamine (Gibco 25030-024) - avoid repeat thawing at 37 degrees	2mL
P/S (Gibco, cat 15140-122)	2.2mL
2-Mercaptoethanol (31350-010)	0.22mL
Total	207.42mL
LIF 10ug/mL (1000x, diluted in PBS + 0.1% BSA. Sourced from CSCR, Sally Lees)	1uL per mL of medium

Supplementary Table 3 2i mixture recipe

2i Mixture	
Components	**Volume**
CHIR99021 (Stemgent, 04-0004 - make 10mM solution by adding 428.9uL of DMSO)	0.3uL per mL of ESM (final conc. 3uM)
PD0325901 (Stemgent, 04-0006 - make 10mM solution by adding 414.8uL of DMSO)	0.1uL per mL of ESM (final conc. 1uM)

Supplementary Table 4 cell culture dish size and volume

Dish	Media	Accutase	Area (cm2)	Diameter cm (actual)
96-well plate	200 uL/well	30-50 uL	0.3	0.6
24-well plate	1 mL/well	200 uL	1.8	1.5
12-well plate	2 mL/well	300uL	3.5	2.1
6-well plate	3-4 mL/well	400 uL	9.6	3.5
6-cm dish	6 mL	0.6 mL	21.2	5.2
10-cm dish	10 mL	1 mL	60	8.7
15-cm dish	30 mL	3 mL	154	14

Supplementary Table 5 cell seeding densities

Cell Line	Media	Dish type	Number of cells seeded	Cells per mL
B6xSD7	Serum + 2i/LIF	12-well plate	200,000	100,000
	N2B27 + 2i/LIF	12-well plate	200,000	100,000
V6.5	Serum + 2i/LIF	12-well plate	100,000	50,000
	N2B27 + 2i/LIF	12-well plate	100,000	50,000
TET TKO mTET1-CD	Serum + 2i/LIF	10 cm dish	2,750,000	275,000
	N2B27 + 2i/LIF	10 cm dish	2,750,000	275,000

Supplementary Table 6 freezing media recipe

Freezing media recipe	
Component	**Volume**
FBS-based ESM (without 2i)	37.5 mL
FBS (15% final concentration)	7.5 mL
DMSO (10% final concentration)	5 mL
Total	50 mL

Supplementary Table 7 forward priming index sequences

Forward priming index sequences	
Primer	AATGATACGGCGACCACCGAGATCTACACNNNNNNNNACACTCTTTCCCTACACGACGCTCTTCCGATCT
Index	
FPI 06	ACATTGGC
FPI 08	CATCAAGT

Supplementary Table 8 reverse priming index sequences

Reverse priming index sequences							
Primer	CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT						
Reverse priming index							
01	AACGTGAT	25	AGTCACTA	49	GATAGACA	73	AATGTTGC
02	AAACATCG	26	ATCCTGTA	50	GCCACATA	74	ACACGACC
03	ATGCCTAA	27	ATTGAGGA	51	GCGAGTAA	75	ACAGATTC
04	AGTGGTCA	28	CAACCACA	52	GCTAACGA	76	AGATGTAC
05	ACCACTGT	29	CAAGACTA	53	GCTCGGTA	77	AGCACCTC
06	ACATTGGC	30	CAATGGAA	54	GGAGAACA	78	AGCCATGC
07	CAGATCTG	31	CACTTCGA	55	GGTGCGAA	79	AGGCTAAC
08	CATCAAGT	32	CAGCGTTA	56	GTACGCAA	80	ATAGCGAC
09	CGCTGATC	33	CATACCAA	57	GTCGTAGA	81	ATCATTCC
10	ACAAGCTA	34	CCAGTTCA	58	GTCTGTCA	82	ATTGGCTC
11	CTGTAGCC	35	CCGAAGTA	59	GTGTTCTA	83	CAAGGAGC
12	AGTACAAG	36	CCGTGAGA	60	TAGGATGA	84	CACCTTAC
13	AACAACCA	37	CCTCCTGA	61	TATCAGCA	85	CCATCCTC
14	AACCGAGA	38	CGAACTTA	62	TCCGTCTA	86	CCGACAAC
15	AACGCTTA	39	CGACTGGA	63	TCTTCACA	87	CCTAATCC
16	AAGACGGA	40	CGCATACA	64	TGAAGAGA	88	CCTCTATC
17	AAGGTACA	41	CTCAATGA	65	TGGAACAA	89	CGACACAC
18	ACACAGAA	42	CTGAGCCA	66	TGGCTTCA	90	CGGATTGC

19	ACAGCAGA	43	CTGGCATA	67	TGGTGGTA	91	CTAAGGTC
20	ACCTCCAA	44	GAATCTGA	68	TTCACGCA	92	GAACAGGC
21	ACGCTCGA	45	GACTAGTA	69	AACTCACC	93	GACAGTGC
22	ACGTATCA	46	GAGCTGAA	70	AAGAGATC	94	GAGTTAGC
23	ACTATGCA	47	AGTCACTA	71	AAGGACAC	95	GATGAATC
24	AGAGTCAA	48	ATCCTGTA	72	AATCCGTC	96	GCCAAGAC